AFRL-IF-RS-TR-2006-295
Final Technical Report
September 2006

VERY LARGE-SCALE MULTIUSER DETECTION (VLSMUD)

Princeton University

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

STINFO FINAL REPORT

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

NOTICE AND SIGNATURE PAGE

AFRL-IF-RS-TR-2006-295 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/s/ /s/

DAVID HENCH
Work Unit Manager

WARREN H. DEBANY, Jr.
Technical Advisor, Information Grid Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
SEP 2006

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
Apr 05 – Aug 06

4. TITLE AND SUBTITLE
VERY LARGE SCALE MULTI-USER DETECTION (VLMUD)

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA8750-05-2-0192

5c. PROGRAM ELEMENT NUMBER
35885G

6. AUTHOR(S)
H. Vincent Poor

5d. PROJECT NUMBER
VLSM

5e. TASK NUMBER
UD

5f. WORK UNIT NUMBER
01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Princeton University
Office of Research and Project Administration
New South Building, Princeton NJ 08544-0036

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFRL/IFGC
525 Brooks Rd
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-IF-RS-TR-2006-295

12. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA# 06-680

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This document is the final report for award FA8750-05-2-0192. It summarizes findings in our study of Very Large Scale Multi-user Detection. This work has involved two distinct efforts: the first addressed the analysis of iterative channel estimation and multi-user detection in multi-path DS-CDMA channels, and the second addressed the issue of energy efficiency in multi-hop CDMA networks.

15. SUBJECT TERMS
Iterative channel estimation, multi-path DS-CDMA channels, multi-hop CDMA networks, multi-user detection

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	UL	45	David Hench
					19b. TELEPHONE NUMBER (*Include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

1. SUMMARY

Our work in this area has involved two distinct efforts: the first addresses the analysis of iterative channel estimation and multiuser detection in multipath DS-CDMA channels [1], and the second addresses the issue of energy efficiency in multi-hop CDMA networks [2].

2. INTRODUCTION

Multiuser detection refers to the data demodulation in multiple-access communication networks. It is a key technology in the development of high-capacity wireless communications and sensor networks. This study was concerned with multiuser detection in networks with very large numbers of nodes, i.e., very large scale multiuser detection. Our work has focused on two aspects of this problem. One of these is in establishing effective methods of analysis for multiuser receivers that iterate between multiuser detection (and channel decoding) and channel estimation [1]. Channel estimation is necessary in wireless receivers because of the dynamic fading that characterizes wireless communication channels. Although such iterative receivers have been proposed before, their analysis (and thus their effectiveness) has been an open question. The second aspect of this problem is that of the effects of multiuser detection on the energy efficiency of multi-hop networks [2], which are of considerable interest in a variety of applications of interest to the Air Force. Energy efficiency is often a primary consideration in design and deployment, and multiuser detection can have considerable effects on this efficiency. The purpose of this work is to develop methods that can be used to characterize these effects in support of design criteria for the such networks.

3. METHODS, ASSUMPTIONS AND PROCEDURES

This work is described in detail in the two papers [1, 2], copies of which are attached as Appendices. Generally speaking the primary assumption is that of a direct-sequence code-division multiple-access (DS-CDMA) network with a very large number of nodes. The principal analytical method used to examine the very large-scale case, is the large-system limit of DSCDMA systems, in which the spreading gain and the number of nodes both increase without bound while their ratio (the “system load”) remains constant. This analytical device allows for closed-form analysis in both of the problems addressed. To analyze the iterative channel estimation and multiuser detection receiver in [1] the main additional techniques are those used to study the convergence of iterative algorithm. To study the energy efficiency of multi-hop networks in [2] a game-theoretic approach is used, in which the terminals in the network work are view as “economic” agents competing for radio resources to transmit their messages as efficiently as possible (i.e.,, with maximal bits transmitted successfully per joule of battery energy).

4. RESULTS AND DISCUSSION

The results are described in detail in the accompanying publications [1, 2]. In the problem of iterative channel estimation and multiuser detection of [1], the results are in the form of characterization of the conditions (in terms of system load and SNR) under which the iterative methods converge, and the resulting performance. In the problem of energy efficiency in multi-hop networks of [2], the results are in the form of the energy efficiency per node (bits-per-joule) as a function of the type of multiuser detector used by the terminals.

5. CONCLUSIONS

For detailed conclusions, please see the accompanying publications, [1,2], in the appendix. Briefly, it is seen in [1] that iteration can substantially improve upon the traditional (noniterative) methods of multiuser receiver implementation. In [2], it is seen that the use of multiuser detection can be a major determinant in the energy efficiency of multi-hop multiuser systems.

APPENDIXES

[1] H. Li, S. Betz and H. V. Poor, "Performance Analysis of Iterative Channel Estimation and Multiuser Detection in Multipath DS-CDMA Channels," *IEEE Transactions on Signal Processing*, accepted for publication. [Originally submitted December 14, 2005; revision submitted March 13, 2006; accepted August 15, 2006.]

[2] S. Betz and H. V. Poor, "Energy Efficiency in Multi-hop CDMA Networks: A Game Theoretic Analysis," (invited paper). *Proceedings of the Workshop on Multi-Layer Modelling and Design of Multi-Hop Wireless Networks (MLMD'06),* Minneapolis, MN, July 12 - 15, 2006.

Appendix A

Performance Analysis of Iterative Channel Estimation and Multiuser Detection in Multipath DS-CDMA Channels

Husheng Li, Sharon M. Betz and H. Vincent Poor

Abstract

This paper examines the performance of decision feedback based iterative channel estimation and multiuser detection in channel coded aperiodic DS-CDMA systems operating over multipath fading channels. First, explicit expressions describing the performance of channel estimation and parallel interference cancellation based multiuser detection are developed. These results are then combined to characterize the evolution of the performance of a system that iterates among channel estimation, multiuser detection and channel decoding. Sufficient conditions for convergence of this system to a unique fixed point are developed.

I. Introduction

Direct sequence code division multiple-access (DS-CDMA) has been selected as the fundamental signaling technique for third generation (3G) wireless communication systems, due to its advantages of soft user capacity limit and inherent frequency diversity. However, it suffers from multiple-access interference (MAI) caused by the non-orthogonality of spreading codes, particularly for heavily loaded systems. Therefore, techniques for mitigating the MAI, namely multiuser detection, have been the subject of an intensive research effort over the past two decades. It is well known that multiuser detection can substantially suppress MAI, thus improving system performance. Maximum likelihood (ML) multiuser detection [28] was proposed in the early 1980s, and achieves the optimal performance at the cost of prohibitive computational

Husheng Li is with Qualcomm Inc., San Diego, CA, 92121, USA (email: hushengl@qualcomm.com). Sharon M. Betz is with Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA (email: sbetz@princeton.edu). H. Vincent Poor is with Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA (email: poor@princeton.edu). This research was supported by the Air Force Research Laboratory under Cooperative Agreement No. FA8750-05-2-01 92.

cost when the number of users is large. For practical implementation, suboptimal algorithms, such as the linear minimum mean square error (LMMSE) detector [21] or decorrelator [29], allow a tradeoff between complexity and performance. It should be noted that the technique of multiuser detection is being applied in existing CDMA systems, such as EV-DO Revision A systems [12].

In recent years, the turbo principle, namely the iterative exchange of soft information among different blocks in a communication system to improve the system performance, has been applied to combine multiuser detection with channel decoding [1][22][24][26][27][31]. In such turbo multiuser detectors, the outputs of channel decoders are fed back to the multiuser detector, thus enhancing the performance iteratively. Turbo multiuser detection based on the maximum *a posteriori* probability (MAP) detection and decoding criterion has been proposed in [30][31] together with a lower complexity technique based on interference cancellation and LMMSE filtering. Further simplification is obtained by applying parallel interference cancellation (PIC) [1] for multiuser detection, where the decisions of the decoders are directly subtracted from the original signal to cancel the MAI.

Practical wireless communication systems usually experience fading channels, whose state information is unknown to the receiver. Thus practical systems need to consider detection and decoding with uncertain channel state information. In the context of short code CDMA systems, blind multiuser detection can be accomplished without explicit channel estimation by using subspace and other techniques [32]. An alternative receiver structure adopts an explicit channel estimation block and carries out the decoding with the corresponding channel estimate. In systems without decision feedback, the channel estimation block is cascaded with the decoder and operates as a front end for the subsequent blocks. With such a receiver structure, the channel estimates can be obtained with training symbols [6] or with blind estimation algorithms [33]. Explicit expressions for the performance of such channel estimation schemes are given in [17] and the corresponding impact on multiuser detection is discussed in the large system limit in [9] and [18]. In systems with decision feedback, the decisions of the decoder are fed back to the channel estimator to enhance its performance. In such systems, the channel estimator and the decoder can operate either simultaneously [25] or successively [7] [13] [23]. An example of the former strategy applied to ML sequence detection in uncertain environments is proposed in [25]; called per-survivor processing, tentative decisions are immediately fed back to the channel estimation algorithm and the corresponding estimates are used for the detection of future symbols. In the latter strategy, the decisions are fed back only when the entire current decoding procedure is finished. For example, in [13], an expectation maximization (EM) channel estimation algorithm, combined with successive interference cancellation, is proposed. Joint channel estimation and data detection algorithms for uncoded single-antenna and multiple-antenna systems are discussed in [8] and [7], respectively. In channel coded systems, iteration can achieve better performance when the turbo principle is applied, due to the redundancy introduced by the code structure. In [23], an iterative algorithm is proposed and analyzed for channel estimation and decoding of low-density parity-check (LDPC) coded quadrature amplitude modulation (QAM) systems.

In this paper, we consider channel-coded CDMA systems operating over multipath fading channels whose channel state information is unknown to the receiver. To demodulate and decode such systems, we apply the turbo principle to both channel estimation and multiuser detection. As shown in Figure 1, we consider a receiver that feeds back decisions from channel decoders to both an ML channel estimator and a PIC multiuser detector. The iteration is initialized with training symbol based channel estimation and a non-iterative multiuser detection. The receiver structure is similar to those proposed in [2][15][20]. However, this paper is focused mainly on the performance analysis of such structures using semi-analytic methods. We analyze the contributions to the variance of the channel estimation error due to noise and decision feedback error, and the variance of the residual MAI after PIC. We then use this analysis to describe the decoding process as an iterative mapping. We also propose conditions assuring convergence of this iterative mapping to a unique fixed point. We further compute the asymptotic multiuser efficiency (AME) [29] of this overall system, under some mild assumptions on the channel decoders. It should be noted that the analysis in this paper is based on large sample and large system analysis.

The remainder of this paper is organized as follows. Section II introduces the signal model and the channel decoder used in our analysis. The performance analyses of ML channel estimation and PIC multiuser detection are given in Section III and Section IV, respectively. Based on these results, the corresponding iterative mapping is described and analyzed in Section V. Numerical results and conclusions are given in Section VI and Section VII, respectively. The notations used in this paper are explained as follows.

- Throughout this paper, if no special note is given, we denote vectors with small letters in bold fonts, matrices with capital letters in bold fonts and scalars with non-bold fonts.
- For any variable X, we denote the corresponding estimate from the decision feedback by $\hat{X}$ and the corresponding error $X - \hat{X}$ by δX.

- Superscript T denotes transposition and superscript H denotes conjugate transposition.
- $\mathbf{I}$ denotes the identity matrix.
- $\lceil x \rceil$ denotes the smallest integer larger than or equal to x.
- $\mathrm{mod}(i, j)$ denotes the modulo of i with respect to j, with the convention of $\mathrm{mod}(i, i) = i$.
- For a matrix $\mathbf{A}_{m\times n}$, $\|\mathbf{A}\|_F \triangleq \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}^2}$ is the Frobenius norm of $\mathbf{A}$.

II. SIGNAL MODEL

A. Signal Model

We consider a synchronous uplink long code (aperiodic) DS-CDMA system, with identical channel coding, binary phase-shift keying (BPSK) modulation, K active users, spreading gain N, system load $\beta = \frac{K}{N}$, and identical transmission rates for all users. The transmitted symbols experience multipath fading. We adopt a block fading model and denote by M the coherence time, measured in the number of symbol periods, over which the channel is stationary. Within a coherence period, the chip matched filter output of the receiver at symbol period t can be collected into an N-vector given by

$$\mathbf{r}(t) = \sum_{k=1}^{K} b_k(t) \sum_{l=1}^{L} a_{kl}\mathbf{s}_{kl}(t) + \mathbf{n}(t), \qquad t = 1, 2, ...M, \tag{1}$$

where L denotes the number of resolvable paths per user, $b_k(t)$ denotes the channel coded binary symbols, a_{kl} denotes the channel gain of the l-th path of user k, $\mathbf{s}_{kl}(t)$ denotes the binary spreading code with $\|\mathbf{s}_{kl}(t)\| = 1$ received from user k along path l at time t and $\mathbf{n}(t)$ is an N-vector of independent and identical distributed (i.i.d.) circularly symmetric complex Gaussian (CSCG) [1]noise variables with (normalized) variance σ_n^2. It should be noted that although the assumption of synchronicity is valid in time division duplexing (TDD) systems, it does not hold for many frequency division duplexing (FDD) systems. However, as it will be shown, the results from the analysis of synchronous systems are also reasonably valid, though not exactly the same, in the case of asynchronous systems.

For the system model, we have the following assumptions.

Assumption II.1: The channel gains $\{a_{kl}\}$ are independently CSCG distributed with zero means and variances $\frac{1}{L}$. We consider only the case of large L, which implies that $\sum_{l=1}^{L}|a_{kl}|^2 \approx 1$, $k = 1, ..., K$; thus all users achieve the same performance with maximal ratio combining (MRC).

[1]A complex random variable is CSCG distributed if its real and imaginary parts are mutually independent Gaussian random variables with zero mean and identical variance.

Assumption II.2: We ignore intersymbol interference (ISI) and assume that the spreading codes received along different paths of a given user are mutually independent (*independent model*).

Assumption II.3: Based on Assumption II.2, the crosscorrelations $\rho_{klmn}(t) \triangleq \mathbf{s}_{kl}(t)^T \mathbf{s}_{mn}(t)$ (note that $\rho_{klkl}(t) = 1$) satisfy

- $E\{\rho_{klmn}(t)\} = 0$, if $(k, l) \neq (m, n)$;
- $E\{\rho^2_{klmn}(t)\} = \frac{1}{N}$, if $(k, l) \neq (m, n)$;
- $E\{\rho_{klmn}(t)\rho_{pqrs}(t)\} = 0$, if $(k, l, m, n) \neq (p, q, r, s)$.

The above assumptions simplify the performance analysis substantially. Moreover, these assumptions are reasonable for practical systems due to the following reasons:

- Assumption II.1 is based on the fact that more propagation paths are resolvable in CDMA systems than narrow band systems, particularly in environments with abundant scattering (e.g., indoor environment). With this assumption, we ignore the impact of the fluctuation of received power incurred by the multipath fading, and consider only the impairment caused by the channel estimation error.
- Assumption II.2 is unrealistic since these sequences are shifted versions of each other (*shifted model*). However, the accuracy of the results dependent upon this assumption is validated with numerical results in Section VI and asymptotic analysis given in Appendix 1.

B. Receiver Structure

The structure of receiver is shown in Figure 1. The channel coefficients are estimated in the channel estimator, which operates in a 'semi-blind' way. Training symbols are available to obtain an initial estimate in the first iteration. In the further iterations the information symbol decisions from channel decoders are assumed to be correct. Then, both the training symbols and fed back decisions are considered as training symbols and used for ML channel estimation. A multiuser detector is used to mitigate the MAI and its outputs are de-interleaved and decoded in the channel decoder. In the multiuser detector, we use the LMMSE algorithm in the first iteration and the PIC algorithm with the aid of hard decision feedback in the succeeding iterations. We follow the standard procedure in turbo multiuser detection [1][13][22][30] to reconstruct the channel symbols from the channel decoder output. Then these channel symbol estimates are interleaved and fed back to the multiuser detector and channel estimator to enhance the performance iteratively.

We denote by $\hat{b}_k(t)$ the estimated binary channel symbol of user k at symbol period t that is fed back from the channel decoder. For simplicity, we use hard decision feedback and denote the feedback symbol error rate by P_e. The decision feedback error is denoted by $\delta b_k(t) \triangleq b_k(t) - \hat{b}_k(t)$. Supposing that both $b_k(t)$ and $\delta b_k(t)$ are symmetrically distributed, it is easy to check that

- $E\left\{\delta b_k(t)\right\} = 0$;
- $E\left\{b_k(t)\delta b_k(t)\right\} = 2P_e$;
- $E\left\{\delta b_k^2(t)\right\} = 4P_e$.
- $E\left\{\delta b_k(m)\delta b_l(n)\right\} = 0$, when $(k, m) \neq (l, n)$.

It should be noted that, in practical systems, soft decision feedback will achieve better performance than hard decision feedback. However, the performance of channel estimation with soft decision feedback is determined by both the first and second moments of the decision feedback error [17]. Thus the corresponding analysis of performance evolution is more complicated than the case of hard decision feedback. Therefore, we adopt hard decision feedback in order to simplify the system performance analysis.

For the decision feedback from channel decoders, we have the following reasonable assumption, which simplifies the analysis and is also used in [1].

Assumption II.4: The codeword length is assumed to be large enough so that the transmitted symbols are coded over many coherence periods. The decision feedbacks $\left\{\hat{b}_k(t)\right\}$ are mutually independent for different k or t.

III. Performance Analysis of Channel Estimation

In this section, we discuss the performance of channel estimation. First, we explain the training symbol based ML channel estimation algorithm that is used in the first iteration. Then, we consider the estimation of the channel coefficients with only hard decision feedback from the channel decoders. Finally, we extend the performance results to channel estimation with both training symbols and decision feedback, the latter of which is used in the further iterations.

In applying the turbo principle, to avoid the reuse of information, only observations $\{\mathbf{r}(t)\}_{t \neq i}$ are used in the channel estimation for multiuser detection in symbol period i. Thus the corresponding channel estimation error is independent of $\mathbf{r}(i)$. However, for simplicity of discussion, we still assume that all M received signals are used for the channel estimation while retaining this independence assumption. For large M, this results in only a small error in the analysis.

In the following discussion of channel estimation and PIC, we regard the channel gains $\{a_{kl}\}$ and the spreading codes $\{\mathbf{s}_{kl}\}$ as realizations of random variables. Only the transmitted symbols, decision feedback errors and noise are considered as random variables. Throughout this paper, all expectations, denoted as $E\{\cdot\}$, are over the distributions of these three variables. Thus our results are conditioned on the realizations of $\{a_{kl}\}$ and $\{\mathbf{s}_{kl}\}$. However, by the strong law of large numbers, we will see that we can obtain identical results for almost every realization of $\{a_{kl}\}$ and $\{\mathbf{s}_{kl}\}$ in the large system limit ($K, N \to \infty$).

A. Training Symbol Based ML Channel Estimation

First we assume that there are M training symbols, channel symbols known to the receiver, within a single coherence period. For simplicity in deriving the channel estimate, we stack the chip matched filter output of the signal corresponding to these training symbols, rewriting (1) as

$$\mathbf{r} = \mathbf{S}\mathbf{a} + \mathbf{n}, \tag{2}$$

where

$$\begin{aligned}
\mathbf{r} &= \left(\mathbf{r}^H(1), ..., \mathbf{r}^H(M)\right)^H_{NM\times 1}, \\
\mathbf{n} &= \left(\mathbf{n}^H(1), ..., \mathbf{n}^H(M)\right)^H_{NM\times 1}, \\
\mathbf{a} &= (a_{11}, a_{12}, ..., a_{KL})^T_{KL\times 1}, \\
\mathbf{S} &= \left((\mathbf{S}(1)\mathbf{B}(1))^T, ..., (\mathbf{S}(M)\mathbf{B}(M))^T\right)^T_{NM\times KL}, \\
\mathbf{B}(m) &= \begin{pmatrix} b_1(m)\mathbf{I}_{L\times L} & 0 & \cdots & 0 \\ 0 & b_2(m)\mathbf{I}_{L\times L} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_K(m)\mathbf{I}_{L\times L} \end{pmatrix}_{KL\times KL} \\
\mathbf{S}(m) &= (\mathbf{s}_{11}(m), \mathbf{s}_{12}(m), ..., \mathbf{s}_{KL}(m))_{N\times KL}, \qquad m = 1, ..., M.
\end{aligned}$$

Applying the ML criterion and the normality of the noise, we can obtain the ML channel estimate, which is given by

$$\begin{aligned}
\hat{\mathbf{a}} &= \arg\max_{\mathbf{a}} P(\mathbf{r}|\mathbf{a}) \\
&= \arg\min_{\mathbf{a}} \|\mathbf{r} - \mathbf{S}\mathbf{a}\| \\
&= (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{r} \\
&= \mathbf{R}^{-1}\mathbf{y}, \qquad (3)
\end{aligned}$$

where $\mathbf{R} = \mathbf{S}^T\mathbf{S}$ and $\mathbf{y} = \mathbf{S}^T\mathbf{r}$.

It follows directly that the channel estimation error is

$$\begin{aligned} \delta\mathbf{a} &= \mathbf{a} - \hat{\mathbf{a}} \\ &= -\mathbf{R}^{-1}\mathbf{S}^T\mathbf{n}, \end{aligned}$$

from which it is obvious that this error has zero mean and covariance $\mathbf{\Sigma_a} \triangleq E\left\{\delta\mathbf{a}\delta\mathbf{a}^H\right\} = \sigma_n^2\mathbf{R}^{-1}$.

For a finite M, we can compute trace $\{\mathbf{R}^{-1}\}$ in the large system limit (i.e. when $K, N \to \infty$ while keeping the system load, $\frac{K}{N} = \beta$, constant). For a system with system load β, it is well known that as $K \to \infty$, $\frac{K}{\text{trace}\{\hat{\mathbf{R}}^{-1}\}}$ converges to the multiuser efficiency of a decorrelator, namely $1-\beta$ [29]. $\frac{\mathbf{R}}{M}$ is equivalent to the covariance matrix of a system with equivalent system load $\beta' = \frac{KL}{MN} = \frac{L}{M}\beta$. Thus as $K, N \to \infty$, we have

$$\frac{\text{trace}\left\{\mathbf{\Sigma_a}\right\}}{M} \to \frac{\sigma_n^2}{M - L\beta}.$$

Therefore, for sufficiently large K and N, the variance of channel estimation error is given by

$$\Delta_a = \frac{\sigma_n^2}{M - L\beta}, \tag{4}$$

which can be approximated by $\Delta_a \approx \frac{\sigma_n^2}{M}$ when M is sufficiently large.

It should be noted that, in asynchronous systems, we can remove part of the chips in the first and the last symbol periods to obtain a similar matrix $\mathbf{S}_{NM-d_{max}\times KL}$, where d_{max} denotes the largest time offsets of different users, measured in chips. Since the training symbols have been incorporated into the spreading codes, we can consider the columns of $\mathbf{S}$ as random $(NM - d_{max})-$ vectors, regardless of the time offsets of different users. Therefore, the variance of channel estimation error in asynchronous systems is similar to that of synchronous systems when M is sufficiently large.

B. Channel Estimation with Decision Feedback

1) Algorithm: When decision feedback is used in place of training symbols to derive the 'ML' channel estimates[2], a process that assumes that the decision feedback is free of error, the channel estimation error is caused by both the thermal noise and the decision feedback error. On applying (3), the channel estimate

[2]By 'ML' estimates, we mean using the expression obtained from the training symbol based estimation, but with symbols obtained from decision feedback. It is not an exact ML estimate since the distribution of the decision feedback error is not considered.

with decision feedback is given by

$$\begin{aligned}\hat{\mathbf{a}} &= \hat{\mathbf{R}}^{-1}\hat{\mathbf{y}} \\ &= \hat{\mathbf{R}}^{-1}\hat{\mathbf{S}}^T(\mathbf{S}\mathbf{a}+\mathbf{n}) \\ &= \mathbf{a}+\hat{\mathbf{R}}^{-1}\hat{\mathbf{S}}^T(\delta\mathbf{S}\mathbf{a}+\mathbf{n}),\end{aligned}$$

where $\delta\mathbf{S} \triangleq \mathbf{S}-\hat{\mathbf{S}}$, $\hat{\mathbf{R}} \triangleq \hat{\mathbf{S}}^T\hat{\mathbf{S}}$, $\hat{\mathbf{y}} \triangleq \hat{\mathbf{S}}^T\mathbf{r}$ and $\hat{\mathbf{S}}$ is the version of $\mathbf{S}$ in (3) obtained from the decision feedback, which is given by

$$\begin{aligned}\hat{\mathbf{S}} &= \left(\left(\mathbf{S}(1)\hat{\mathbf{B}}(1)\right)^T, ..., \left(\mathbf{S}(M)\hat{\mathbf{B}}(M)\right)^T\right)^T_{NM\times KL}, \\ \hat{\mathbf{B}}(m) &= \begin{pmatrix} \hat{b}_1(m)\mathbf{I}_{L\times L} & 0 & \cdots & 0 \\ 0 & \hat{b}_2(m)\mathbf{I}_{L\times L} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{b}_K(m)\mathbf{I}_{L\times L} \end{pmatrix}_{KL\times KL}.\end{aligned}$$

Hence, the channel estimation error can be decomposed into two parts

$$\begin{aligned}\delta\mathbf{a} &= -\hat{\mathbf{R}}^{-1}\hat{\mathbf{S}}^T(\delta\mathbf{S}\mathbf{a}+\mathbf{n}) \\ &= \delta\mathbf{a}_f+\delta\mathbf{a}_n, \end{aligned} \tag{5}$$

where $\delta\mathbf{a}_f \triangleq -\hat{\mathbf{R}}^{-1}\hat{\mathbf{S}}^T\delta\mathbf{S}\mathbf{a}$ and $\delta\mathbf{a}_n \triangleq -\hat{\mathbf{R}}^{-1}\hat{\mathbf{S}}^T\mathbf{n}$ denote the channel estimation error due to the decision feedback error and the thermal noise, respectively. It is reasonable to assume that $\delta\mathbf{a}_f$ and $\delta\mathbf{a}_n$ are mutually independent. (Recall our assumption concerning the use of only measurements $t \neq i$ in estimating gains at time i.)

It is difficult to tackle the calculation of $\delta\mathbf{a}$ due to the matrix inversion $\hat{\mathbf{R}}^{-1}$. However, we can approximate $\hat{\mathbf{R}}^{-1}$ by $\frac{\mathbf{I}_{KL\times KL}}{M}$ when P_e is sufficiently small. This approximation is justified by the following lemma.

Lemma III.1: When fixing K and N, we have

$$M\hat{\mathbf{R}}^{-1} \rightarrow \mathbf{I}_{KL\times KL},$$

almost surely [3] as $M \rightarrow \infty$ and $P_e \rightarrow 0$.

Proof: According to the definition of $\hat{\mathbf{R}}$, we have

$$\hat{\mathbf{R}}^{-1} = \mathbf{R}^{-1}+\mathbf{R}^{-1}\mathbf{A},$$

[3]Here, a matrix is considered as a point in the probability space and the metric is induced by a matrix norm.

where $\mathbf{A} = \left(\mathbf{I} - \delta\mathbf{R}\mathbf{R}^{-1}\right)^{-1} - \mathbf{I}$. According to the error analysis of matrix inversion in [11], we have[4]

$$E\left\{\|\mathbf{A}\|_F\right\} \leq E\left\{\frac{\|\delta\mathbf{R}\mathbf{R}^{-1}\|_F}{1 - \|\delta\mathbf{R}\mathbf{R}^{-1}\|_F}\right\} = O(P_e),$$

which tends to 0 as $P_e \to 0$. Thus, we have

$$E\left\{\left\|\hat{\mathbf{R}}^{-1} - \mathbf{R}^{-1}\right\|_F\right\} \leq \left\|\mathbf{R}^{-1}\right\|_F E\left\{\|\mathbf{A}\|_F\right\} \to 0,$$

as $P_e \to 0$. Therefore, $\hat{\mathbf{R}}^{-1}$ converges to $\mathbf{R}^{-1}$ almost surely as $P_e \to 0$.

Applying the strong law of large numbers and the fact that the diagonal elements in

$$\mathbf{R} = \sum_{m=1}^{M} (\hat{\mathbf{B}}(m)\mathbf{S}(m))^T \mathbf{S}(m)\hat{\mathbf{B}}(m)$$

are M and the off-diagonal elements in $(\hat{\mathbf{B}}(m)\mathbf{S}(m))^T \mathbf{S}(m)\hat{\mathbf{B}}(m)$ are independent for different values of m and have zero mean, we obtain that, while keeping K and N fixed, $\frac{\mathbf{R}}{M} \to \mathbf{I}_{KL\times KL}$ almost surely, as $M \to \infty$. Since the elements of $\mathbf{R}^{-1}$ are continuous functions of those in $\mathbf{R}$ in a neighborhood of $\mathbf{R} = M\mathbf{I}_{KL\times KL}$, we also have $M\mathbf{R}^{-1} \to \mathbf{I}_{KL\times KL}$ as $M \to \infty$. This completes the proof. ■

Therefore, we can further approximate $\hat{\mathbf{R}}^{-1}$ by $\frac{\mathbf{I}_{KL\times KL}}{M}$ for large M and small P_e. For simplicity, our further discussion of $\delta\mathbf{a}_f$ will be based on this approximation, which will be validated by numerical results. Consequently, in the following discussions, we use the approximations

$$\delta\mathbf{a}_f = -\frac{1}{M}\hat{\mathbf{S}}^T \delta\mathbf{S}\mathbf{a},$$

and

$$\delta\mathbf{a}_n = -\frac{1}{M}\hat{\mathbf{S}}^T \mathbf{n}.$$

2) Covariance matrix of channel estimation error: We denote the covariance matrices of $\delta\mathbf{a}$, $\delta\mathbf{a}_f$ and $\delta\mathbf{a}_n$ by $\mathbf{\Sigma_a}$, $\mathbf{\Sigma_f}$ and $\mathbf{\Sigma_n}$, respectively, which satisfy $\mathbf{\Sigma_a} = \mathbf{\Sigma_f} + \mathbf{\Sigma_n}$. We first consider the channel estimation error incurred by decision feedback errors. The following lemma shows that the channel estimation error $\delta\mathbf{a}_f$ is asymptotically biased. The proof is given in Appendix II.

Lemma III.2: When keeping K and N fixed, we have

$$E\{\delta\mathbf{a}_f\} \to 2P_e\mathbf{a}, \tag{6}$$

almost surely, as $M \to \infty$.

[4] $x = O(P_e)$ means $\frac{x}{P_e} < \infty$ as $P_e \to 0$.

It should be noted that this bias cannot be removed *a priori* in the estimator since it is dependent on the channel gain, $\mathbf{a}$. However, this bias vanishes as $P_e \to 0$.

An asymptotic expression for the elements in $\mathbf{\Sigma_f}$ is given in the following proposition, whose proof is given in Appendix III, where we also explain that the conclusion also applies to asynchronous case when P_e is sufficiently small.

Proposition III.3: For all i and j, when fixing K and N, we have that (recall that a_{kl} is the channel gain of use k and path l)

$$M \times (\mathbf{\Sigma_f})_{ij} \quad \to \quad \begin{cases} 4P_e \left(|a_{\lceil \frac{i}{L} \rceil,\ mod(i,L)}|^2 + \frac{1}{N} \sum_{k=1,\ k \neq i}^{KL} |a_{\lceil \frac{k}{L} \rceil,\ mod(k,L)}|^2 \right), & \text{if } i = j, \\ 4P_e \left(1 + \frac{1}{N}\right) a_{\lceil \frac{i}{L} \rceil,\ mod(i,L)} a^*_{\lceil \frac{j}{L} \rceil,\ mod(j,L)}, & \text{if } i \neq j \text{ and } \lceil \frac{i}{L} \rceil = \lceil \frac{j}{L} \rceil, \\ 4P_e^2 \left(1 + \frac{1}{N}\right) a_{\lceil \frac{i}{L} \rceil,\ mod(i,L)} a^*_{\lceil \frac{j}{L} \rceil,\ mod(j,L)}, & \text{if } \lceil \frac{i}{L} \rceil \neq \lceil \frac{j}{L} \rceil \end{cases}, \tag{7}$$

almost surely, as $M \to \infty$.

For $\delta \mathbf{a}_n$, which is caused by thermal noise, the corresponding analysis is identical to that of training symbol based estimation. Then, we have

$$\begin{aligned} M\mathbf{\Sigma_n} &= M\text{cov}\left(\hat{\mathbf{R}}^{-1}\hat{\mathbf{S}}^T\mathbf{n}\right) \\ &= M\sigma_n^2\hat{\mathbf{R}}^{-1} \\ &\to \sigma_n^2\mathbf{I}_{KL\times KL}, \end{aligned} \tag{8}$$

almost surely, as $M \to \infty$. Then the covariance matrix of channel estimation error $\mathbf{\Sigma_a} \triangleq E\left\{\delta\mathbf{a}\delta\mathbf{a}^H\right\} = \mathbf{\Sigma_f} + \mathbf{\Sigma_n}$ can be obtained from (7) and (8).

3) Variance of channel estimation error: The variance of channel estimation error can be obtained as a corollary of the previous subsection.

Corollary III.4: On defining $\Delta_a \triangleq \frac{1}{KL}\text{trace}\left\{\mathbf{\Sigma_a}\right\}$, we have

$$M\Delta_a \quad \to \quad \frac{4P_e(1+\beta L)}{L} + \sigma_n^2, \tag{9}$$

almost surely, as $K, N, M \to \infty$.

Thus, when K, N, M are sufficiently large, we have the following approximation

$$\Delta_a \quad \approx \quad \frac{4P_e(1+\beta L)}{LM} + \frac{\sigma_n^2}{M}. \tag{10}$$

It should be noted that the channel estimation error cannot be removed by increasing M although the variance vanishes as $M \to \infty$, since the estimate is biased and the bias cannot be removed *a priori*.

C. Estimation with Both Training Symbols and Decision Feedback

We denote the number of training symbols by M_t and the corresponding percentage by $\alpha = \frac{M_t}{M}$. When the training symbols and decision feedback are combined for channel estimation, the performance is determined by (10), with P_e replaced by $(1-\alpha)P_e$. Decision feedback should only be used along with the training symbols if the resulting variance is smaller than that obtained when only the training symbols are used. Then it is easy to check that, when M and M_t are sufficiently large, $P_{e\,\max}$, the maximum P_e assuring performance improvement when decision feedback is used, is determined by

$$\frac{4(1-\alpha)P_e(1+\beta L)}{LM} + \frac{\sigma_n^2}{M} \leq \frac{\sigma_n^2}{M_t},$$

which results in

$$P_{e\,\max} = \frac{\sigma_n^2 L}{4\alpha(1+\beta L)}, \tag{11}$$

from which we observe that $P_{e\,\max}$ decreases with α and β while increasing with σ_n^2 and L.

IV. PIC AND CHANNEL DECODER

A. Performance Analysis of PIC

For convenience of analysis, the performance of PIC is analyzed based on matched filter (MF) outputs. We drop the index of the symbol period for notational simplicity throughout this section. For a given symbol period, the MF outputs, which form sufficient statistics for multiuser detection, are given by

$$\mathbf{y} = \mathbf{S}^T \mathbf{r}.$$

In PIC based multiuser detection, the MAI reconstructed from the channel estimates and the decoder output is subtracted directly from the MF output of the desired user. Without loss of generality, we take the l-th path of user 1 as an example; then the MF output after PIC, which is contaminated by residual MAI and thermal noise $n_{1l} = \mathbf{s}_{1l}^T \mathbf{n}$, is given by

$$y_{1l} = a_{1l} b_1 + \sum_{m \neq l} a_{1m} \rho_{1l1m} b_1 + I_{1l}, \tag{12}$$

where

$$I_{1l} = \sum_{k=2}^{K} \sum_{m=1}^{L} \rho_{1lkm} \left(a_{km} b_k - \hat{a}_{km} \hat{b}_k \right) + n_{1l},$$

which is the sum of the residual interference and the thermal noise. It is obvious that $E\{I_{1l}\} = 0$. And the corresponding variance is given by

$$\begin{aligned}
\sigma_I^2 \triangleq E\left\{|I_{1l}|^2\right\} &= \frac{1}{N}\sum_{k=2}^{K}\sum_{m=1}^{L} E\left\{|\delta a_{km} b_k + \delta b_k a_{km} - \delta a_{km}\delta b_k|^2\right\} + E\left\{|n_{1l}|^2\right\} \\
&= \frac{1}{N}\sum_{k=2}^{K}\sum_{m=1}^{L} \Big\{ E\left\{|\delta a_{km}|^2\right\} + 4P_e |a_{km}|^2 + 4P_e E\left\{|\delta a_{km}|^2\right\} + 2E\left\{\delta a_{km} a_{km}^*\right\} E\left\{b_k \delta b_k\right\} \\
&\quad -2E\left\{|\delta a_{km}|^2\right\} E\left\{b_k \delta b_k\right\} - 8P_e E\left\{a_{km}\delta a_{km}^*\right\} \Big\} + \sigma_n^2 \\
&\rightarrow \beta L \Delta_a + 4\beta(1-P_e)P_e + \sigma_n^2,
\end{aligned} \tag{13}$$

as $K, N \rightarrow \infty$, where we have applied the fact that $E\left\{|\delta a_{km}|^2\right\} = \Delta_a + 4P_e^2 |a_{km}|^2$, $E\left\{a_{km}\delta a_{km}^*\right\} = 2P_e |a_{km}|^2$, $E\left\{b_k \delta b_k\right\} = 2P_e$. It is easy to check that σ_I^2 is identical for asynchronous systems since different time offsets do not affect the interference power.

It is difficult to apply the central limit theorem to show the asymptotic normality of the PIC output since the variables $\{\delta a_{km}\}$ are mutually correlated across different users and paths. However, numerical results in Section VI will show that the output distribution of PIC can be well approximated by a Gaussian distribution. Thus, in the subsequent sections, we assume that the output of PIC is Gaussian distributed.

According to the properties of the crosscorrelation given in Section II.A, $\rho_{1l1m} \rightarrow 0$ almost surely, as $N \rightarrow \infty$. Thus, for large spreading gain, the interference across different paths of the same user can be ignored. With the normality assumption of the residual MAI, it is easy to show that the variables $\{I_{1l}\}_{l=1,\ldots,L}$ are mutually independent as $N \rightarrow \infty$, which means that channel coded symbol b_1 is transmitted through L independent channels. This assumption simplifies the analysis although it does not hold exactly when N is finite. Thus, we use MRC to collect these L replicas, resulting in the output

$$z_1 = \sum_{l=1}^{L} \hat{a}_{1l}^* a_{1l} b_1 + \sum_{l=1}^{L} \hat{a}_{1l}^* I_{1l}. \tag{14}$$

Applying Lemma III.2, we obtain that, as $M, L \rightarrow \infty$,

$$\begin{aligned}
\sum_{l=1}^{L} \hat{a}_{1l}^* a_{1l} b_1 &= \sum_{l=1}^{L} \left(|a_{1l}|^2 - \delta a_{1l}^* a_{1l}\right) \\
&\rightarrow 1 - \sum_{l=1}^{\infty} E\left\{\delta a_{1l}^*\right\} a_{1l} \\
&= 1 - 2P_e \sum_{l=1}^{\infty} |a_{1l}|^2 \\
&= 1 - 2P_e.
\end{aligned}$$

Moreover, we can obtain that, as $M, L \to \infty$

$$\begin{aligned} E\left\{\left|\sum_{l=1}^{L} \hat{a}_{1l}^* I_{1l}\right|^2\right\} &= \sum_{l=1}^{L} E\left\{|\hat{a}_{1l}^*|^2\right\} \sigma_I^2 \\ &= \left(1 - 2\sum_{l=1}^{L} E\left\{\delta a_{1l}^* a_{1l}\right\} + \sum_{l=1}^{L} E\left\{|\delta a_{1l}^*|^2\right\}\right) \sigma_I^2 \\ &\to \left(1 - 4P_e + 4P_e^2 + L\Delta_a\right) \sigma_I^2 \\ &= ((1 - 2P_e)^2 + L\Delta_a)\sigma_I^2. \end{aligned}$$

Therefore, when M and L are sufficiently large, (14) can be approximated by

$$z_1 \approx (1 - 2P_e)b_1 + n_1, \tag{15}$$

where n_1 is a CSCG random variable with variance of $((1 - 2P_e)^2 + L\Delta_a)\sigma_I^2$. An interesting observation is that the channel estimation error not only increases the interference but also decreases the valid received power of the desired user.

B. Performance of Channel Decoder

At the channel decoder, P_e is a function of the input signal-to-interference-plus-noise ratio (SINR) at the input to the channel decoder given by

$$P_e = g\left(\frac{1}{\text{SINR}}\right), \tag{16}$$

where the function g can be estimated using Monte Carlo simulations. For most practical channel codes, the following assumption is reasonable:

Assumption IV.1: Within a closed interval $\Omega = [0, \sigma_I^{max}]$, function g satisfies

- $g(x)$ monotonically increases with x, and $g(0) = 0$;
- $g(x)$ is continuously differentiable and $g'(0) = 0$.

V. Analysis of System Performance

In this section, we analyze the overall iterative system shown in Figure 1. We consider only the case of small P_e, moderate σ_n^2 and moderate M and note that the analytic results become more precise as P_e and σ_n^2 decrease and M increases. This configuration is reasonable for the decision feedback basedsystems since if M is large, training symbol based channel estimation can be adopted with marginal loss of spectral efficiency; if M is small, it is difficult to carry out coherent detection; and if P_e is large, the iteration diverges. Although the performance analysis of the channel estimation in Section III is based on large M, numerical results in Section VI indicate that expression (10) is still valid for moderate M. We adopt the expressions (10) and (13) in large system limits ($K, N \to \infty$).

A. Iterative Mapping

In this section, we consider the d-th iteration and couple the results from Section III and Section IV to analyze the overall system performance. We can regard the decoding process as an iterative mapping $h : \mathbb{R} \to \mathbb{R}$ in terms of the error probability of the decoder output after the d-th iteration, $P_e^{(d)}$, which is given by (recall that g is defined as the function characterizing the output error probability in terms in input SINR in (16))

$$
\begin{aligned}
P_e^{(d)} &= h(P_e^{(d-1)}) \\
&\approx g\left(D_0 + D_1 P_e^{(d-1)}\right), \qquad (17)
\end{aligned}
$$

where we ignore terms of a smaller order than P_e and $\frac{1}{M}$ since we assume small P_e and large (or moderate) M. Based on (10), (13) and (15), the coefficients D_0 and D_1 are given by

$$
\begin{cases}
D_0 = \sigma_n^2\left(1 + \frac{\beta L}{M} + \frac{L\sigma_n^2}{M}\right) \\
D_1 = 4\left(\beta + \frac{\beta + \sigma_n^2 \beta L^2 + \beta^2 L + \sigma_n^2 L + L\beta\sigma_n^2 + L\left(\sigma_n^2\right)^2}{M}\right)
\end{cases}.
$$

B. Condition for Convergence

A reasonably good initialization, which results in sufficiently small channel estimation error and MAI in the first iteration, is necessary to guarantee the convergence of the iterative mapping described in (17). In the initial stage, only training symbols are used for the channel estimation since no decision feedback is available then. Any non-iterative multiuser detection technique can be applied to the initializing stage. For practical applications, we can use the LMMSE detector, whose performance using imperfect channel estimation can be obtained using the replica method [18].

For convergence, the variance of input interference and noise of the initializing stage, denoted by $\sigma_I^2(0)$ and obtained from the SINR of the LMMSE detector, must satisfy the following conditions:

- $\sigma_I^2(0)$ is located within the interval Ω defined in Section IV.B, namely

$$
\sigma_I^2(0) < \sigma_I^{max}. \qquad (18)
$$

This condition assures a reasonably good initial performance of the iterations.

- The variance of interference and noise decreases with iteration time, namely

$$g(\sigma_I^2(0)) < \frac{\sigma_I^2(0) - D_0}{D_1}. \tag{19}$$

This condition assures that the iterations do not diverge.

C. Condition Assuring the Uniqueness of the Fixed Point

If there exists more than one fixed point, the iteration may become stuck at a suboptimal fixed point and not converge to the optimal one. The following proposition provides a sufficient condition for the uniqueness of the fixed point and the corresponding convergence rate.

Proposition V.1: (1) If there exists a $\gamma < 1$, such that

$$D_1 \leq \frac{\gamma}{\max_{x\in\Omega}\left(g'(x)\right)}, \tag{20}$$

then there exists only one fixed point x_f for the iterative mapping $x_{k+1} = h\left(x_k\right)$, and for every initial point $x_0 \in \Omega$, the mapping converges to x_f with an exponential rate, namely $\|x_k - x_f\| \leq \frac{\gamma^k}{1-\gamma}\|x_0 - x_f\|$.

(2) If there exists an $x_1 \in \Omega$ such that $\frac{1}{g'(x_1)} < D_1 < \frac{x_1}{g(x_1)}$, then there exists a D_0 such that there is more than one fixed point for h.

Proof: (1) The condition $D_1 \leq \frac{\gamma}{\max_{x\in\Omega}(g'(x))}$ implies that $h'(x) = g'(D_0 + D_1 x) \leq \gamma < 1$. Then $h(\cdot)$ is a contraction mapping, and the conclusions follow due to Banach's fixed point theorem [14].

(2) Letting $x_f = g(x_1)$ and setting $D_0 = x_1 - D_1 x_f$, we can show that $D_0 > 0$ due to the assumption that $D_1 < \frac{x_1}{g(x_1)} = \frac{x_1}{x_f}$. It is easy to check that x_f is a fixed point and $g'(D_0 + D_1 x_f) = D_1 g'(x_1) > 1$. Hence, there exists an $\epsilon > 0$ such that for all $x \in (x_f, x_f + \epsilon)$, $g(D_0 + D_1 x) > x$. However, $g(D_0 + D_1 x_2) < x_2$ for $x_2 = g\left(\sigma_I^2(0)\right)$ due to condition (19). If $x_2 < x_f$, there exists at least one fixed point within $(0, x_2)$ since $g(D_0) > 0$; if $x_2 > x_f$, there exists at least one fixed point different from x_f within (x_f, x_2). ■

It should be noted that condition (20) is sufficient but not necessary for the uniqueness of the fixed point. This condition is more stringent than the condition of convergence in (19) since it assures both the uniqueness of the fixed point and the exponential convergence rate. The second part shows that a moderate D_1 may cause multiple fixed points. A useful conclusion drawn from (20) is that this iterative procedure does not work well for those channel codes, such as powerful turbo codes or LDPC codes, that have a steep performance curve (bit error rate versus SINR) which implies a large value of $\max_{x\in\Omega}\left(g'(x)\right)$. This will be demonstrated in numerical simulations in Section VI.

D. Asymptotic Multiuser Efficiency

As is described in [29], the asymptotic multiuser efficiency measures the slope at which the bit-error-rate goes to zero in logarithmic scale, giving intuition into the performance loss from multiuser interference.

Suppose that there is only one fixed point for the iterative mapping h, and let $P_e(\sigma_n^2)$ be this fixed point when the noise power is σ_n^2. Similarly, let $D_0(\sigma_n^2)$ and $D_1(\sigma_n^2)$ be the corresponding values of D_0 and D_1 in (17). It is obvious that $P_e(0) = 0$ and $D_0(0) = 0$.

The asymptotic multiuser efficiency is given by

$$\begin{aligned}
\text{AME} &= \lim_{\sigma_n^2 \to 0} \frac{\sigma_n^2}{D_0(\sigma_n^2) + D_1(\sigma_n^2) P_e(\sigma_n^2)} \\
&= \frac{1}{\frac{dD_0(\sigma_n^2)}{d\sigma_n^2}\Big|_{\sigma_n^2=0} + \frac{d(D_1(\sigma_n^2)P_e(\sigma_n^2))}{d\sigma_n^2}\Big|_{\sigma_n^2=0}}.
\end{aligned}$$

If $H(P_e, \sigma_n^2) = g\left(D_0(\sigma_n^2) + D_1(\sigma_n^2)P_e\right) - P_e$, then $P_e(\sigma_n^2)$ is the unique solution of $H(P_e, \sigma_n^2) = 0$. Applying the assumptions that $g'(0) = 0$ and $P_e(0) = 0$, we have

$$\begin{aligned}
\frac{d(D_1(\sigma_n^2)P_e(\sigma_n^2))}{d\sigma_n^2}\Big|_{\sigma_n^2=0} &= D_1(0)\frac{dP_e(\sigma_n^2)}{d\sigma_n^2}\Big|_{\sigma_n^2=0} \\
&= -D_1(0)\frac{\frac{\partial H(P_e,\sigma_n^2)}{\partial \sigma_n^2}\Big|_{\sigma_n^2=0}}{\frac{\partial H(P_e,\sigma_n^2)}{\partial P_e}\Big|_{\sigma_n^2=0}} \\
&= -D_1(0)\frac{\frac{\partial(D_0(\sigma_n^2)+D_1(\sigma_n^2)P_e)}{\partial \sigma_n^2}\Big|_{\sigma_n^2=0} g'(0)}{D_1(0)g'(0) - 1} \\
&= 0.
\end{aligned}$$

Thus

$$\begin{aligned}
\text{AME} &= \frac{1}{\frac{dD_0(\sigma_n^2)}{d\sigma_n^2}\Big|_{\sigma_n^2=0}} \\
&= \frac{1}{1 + \frac{L\beta}{M}}.
\end{aligned} \tag{21}$$

From (21), we can see that the loss of AME is due to the channel estimation error incurred by the thermal noise. The impact of the decision feedback error vanishes as $\sigma_n^2 \to 0$, while that of the channel estimation error remains.

E. Computational Aspect

The main computational cost of the iterative channel estimation and multiuser detection includes:

- Solving the linear equation $\hat{\mathbf{R}}\hat{\mathbf{a}} = \mathbf{y}$ for ML channel estimation.
- Reconstructing the channel symbols and cancelling the interference.
- Channel decoding.

Since the channel symbol reconstruction is similar to the encoding procedure and the interference cancelation requires only subtractions, this is not a bottleneck of the whole procedure and the corresponding computational cost is of complexity $O(K)$. Real-time channel decoding can also be accomplished in a way similar to Turbo codes. Therefore, the main bottleneck is solving the linear equation for channel estimation.

Direct Gaussian Eliminatation, which is of complexity $O(K^3)$, can be applied to solve the equation $\hat{\mathbf{R}}\hat{\mathbf{a}} = \mathbf{y}$ when K is small. When K is large, iterative techniques of solving linear equations, such as the Jacobi method and the Gauss-Seidel method, can be applied. For assuring the convergence, we cite the following lemma from [10]:

Lemma V.2: The sufficient and necessary condition for the convergence of iterations in solving the linear equation $\mathbf{A}\mathbf{x} = \mathbf{y}$ is that

- $\mathbf{A}$ and $2\,\mathrm{diag}(\mathbf{A}) - \mathbf{A}$ are both positive definite in the Jacobi method[5];
- $\mathbf{A}$ is positive definite in the Gauss-Seidel method.

The Gauss-Seidel method always converges when $\beta < 1$ since $\hat{\mathbf{R}}$ is positive definite when $K < N$. For the Jacobi method, it is easy to check that $\mathrm{diag}(\hat{\mathbf{R}}) = \mathbf{I}_{K\times K}$. Since the largest eigenvalue of $\hat{\mathbf{R}}$ converges to $\left(1+\sqrt{\beta}\right)^2$ [3] almost surely as $K, N \to \infty$, the eigenvalues in $2\,\mathrm{diag}(\hat{\mathbf{R}}) - \hat{\mathbf{R}}$ are less than $2 - \left(1+\sqrt{\beta}\right)^2$ almost surely in the large system limit. Therefore, $\sqrt{\beta} < 1$ is a sufficient condition for the almost sure convergence of Jacobi iteration in the large system limit. Then, when K and N are sufficiently large and $K < N$, we can use either Gauss-Seidel or Jacobi iterations to estimate the channel coefficients efficiently.

VI. Numerical Results

A. Channel Estimation

Figure 2 shows the average variance of the channel estimates versus the coherence time M with the configuration of $\beta = 0.2$, $L = 5$, $M_t = 0$, $P_e = 0.1$ and the signal-to-noise ratio (SNR)= 5dB[6]. The asymptotic results obtained from (10) and the simulation results for finite systems ($N = 100$) with spreading codes for the shifted model are represented by solid and dotted curves, respectively. In this figure, the

[5] $\mathrm{diag}(\mathbf{X})$ denotes a diagonal matrix constituted by the diagonal elements in matrix $\mathbf{X}$

[6] Note that P_e and SNR are not mutually independent; however, we set these two parameters arbitrarily to test the validity of asymptotic results.

estimation error variance caused by decision feedback and noise are denoted by Δ_f and Δ_n, respectively. The corresponding asymptotic results are obtained from the first and the second terms in (10), respectively. We can observe that the asymptotic results match the simulation results well even when M is small. This figure also demonstrates the validity of results based on the independence assumption of the spreading codes given in Section II.A.

B. *Normality of PIC Output*

Figure 3 shows the channel symbol error rate[7] with the configuration of SNR $=$ 10dB, $K = N = 30$ and $P_e = 0.1$ and 0.05. The solid curves represent the results obtained from numerical simulations and the dashed curves represent the results with the assumption that the output of PIC is CSCG distributed. The gap between the numerical results and CSCG based prediction is small, thus justifying the normality assumption of the PIC output.

C. *User Capacity*

We define the user capacity to be the maximum system load β_{max} with which the system can achieve the information bit error rate of 10^{-3}. Two types of channel codes, the convolutional code $(35, 23)_8$ and a turbo code (with two constituent codes $(37, 21)_8$), with bit rate $R = \frac{1}{2}$ and codeword length 1024 are used in this paper and their error rates for both information bits and extrinsic information based channel symbols are shown in Figure 4. The corresponding β_{max}'s for various values of coherence time M, denoted by 'iterative', are given in Figure 5 and Figure 6 for convolutional codes and turbo codes, respectively, with the configuration $\alpha = 0.2$, SNR$=$ 5dB and $L = 5$. The β_{max}'s of the non-iterative LMMSE detector, denoted by 'LMMSE', are given for comparison. We can see that the iterative system achieves substantially higher user capacity than the non-iterative one. The performance of systems with ideal initialization, where actual channel parameters are provided by a genie in the initialization stage, denoted by 'Perfect initialization', implies that a good initialization can improve the performance considerably. Thus, blind or semi-blind non-iterative techniques, which make use of information symbols, can be applied to obtain a better initialization. For comparison, the user capacities of both iterative and non-iterative systems with perfect channel state information are also given in both figures. An interesting observation is that the relative performance gain

[7]This channel symbol error rate is equivalent to bit error rate when the output of PIC is used directly for the detection (without channel decoding).

of iterative systems over the non-iterative ones is smaller for turbo codes than for convolutional codes. This is due to the steeper waterfall region in turbo codes.

VII. Conclusions

In this paper, we have analyzed the performance of decision feedback based iterative channel estimation and multiuser detection in multipath DS-CDMA channels. The decoding process has been described as an iterative mapping in terms of the variance of the channel decoder output, and conditions assuring the convergence and uniqueness of a fixed point have been proposed. Numerical results show that the initialization is important to the iterations, thus necessitating the use of non-iterative blind or semi-blind channel estimation algorithms for initialization purposes. Another observation of interest is that the gain of the iterative process over a non-iterative one is small when a near-optimal channel coding scheme is used.

Appendix I

Validity of Independence Model for Spreading Codes

In (1), for different values of l and m, $\mathbf{s}_{kl}$ and $\mathbf{s}_{km}$ are generated by the same binary sequence with different offsets. Our purpose is to show that if K and N are large enough, we can regard the shifted spreading codes of different paths of a given user as independent sequences. The properties based on this assumption, which are used for the system performance analysis in this paper, include:

- The properties of crosscorrelation ρ_{klmn} in Section II.A.
- The distribution of the eigenvalues of the matrix $\mathbf{S}\mathbf{S}^T$, when developing the expression of Δ_n for finite M and large K in Section III.C. Our assumption means that the corresponding distribution of the shifted model is asymptotically identical to that of the independent model.

It is easy to check the first item using the symmetry of the binary distribution. However, the validity of the second one is non-trivial and is of considerable importance when applying the theory of large random matrices to multipath fading channels. We can tackle this problem by showing that the moments of the eigenvalues in both models are the same via the following lemma.

Lemma I.1: Denote a generic eigenvalue of $\mathbf{S}\mathbf{S}^T$ by λ. Then the m-th moment of λ in the shifted model is given by

$$E\left\{\lambda^m\right\} = \sum_{k=1}^{m} \left(\beta'\right)^k \sum_{m_1+\ldots+m_k=m} c(m_1, \ldots, m_k), \qquad \text{as } K \to \infty,$$

which is the same expression of that of the independent model, and where the definition of $c(m_1, ..., m_k)$ is given in [19] and $\beta' = \frac{LK}{MN}$.

Proof: Using similar arguments to those in [19], we have

$$\begin{aligned} & \frac{1}{N} E\left\{\text{trace}\{(\mathbf{S}\mathbf{S}^T)^m\}\right\} \\ = & \frac{1}{N^{m+1}} \sum_{i_1,\ldots,i_m=1}^{K} \sum_{j_1,\ldots,j_m=1}^{N} E\{V_{i_m,j_1} V_{i_1,j_1} \ldots V_{i_{m-1},j_m} V_{i_m,j_m}\}, \end{aligned} \tag{22}$$

where $V_{i,j} = \sqrt{N}\mathbf{S}_{ij}$.

For any $i_r \neq i_s$, $V_{i_r,j_p} = V_{i_s,j_q}$ when $\lceil \frac{i_r}{L} \rceil = \lceil \frac{i_s}{L} \rceil$ and $j_p - j_q$ equals the offset difference between these two shifted sequences. However, the probability of such events vanishes as $K \to \infty$ since

$$P\left(|i_r - i_s| < L\right) \leq \begin{pmatrix} m \\ 2 \end{pmatrix} \frac{2L+1}{KL} \to 0, \qquad \text{as } K \to \infty.$$

Thus, as $K \to \infty$, the term involving $V_{i,j}$'s of different users, which are mutually independent, dominates the summation in (22). The remaining part of the proof is the same as in [19]. ■

The following lemma (Theorem 30.1 in [5]) provides a sufficient condition for the equality of two probability measures when their moments are identical .

Lemma I.2: Let μ be a probability measure on the real line having finite moments $\alpha_k = \int_{-\infty}^{\infty} x^k \mu(dx)$ of all orders. If the power series $\sum_{k=1}^{\infty} \alpha_k \frac{r^k}{k!}$ has a positive radius of convergence, then μ is the only probability measure with the moments $\{\alpha_m\}_{m=1,2,\ldots}$.

For applying Lemma I.2, we need the following lemma which provides an upper bound for the moments of the eigenvalues.

Lemma I.3: For any eigenvalue λ of $\mathbf{S}\mathbf{S}^T$, there exists a constant $C > \max(1, \beta')$ such that for $m = 1, 2, \ldots$

$$E\left\{\lambda^m\right\} < C^m m^{m-2}. \tag{23}$$

Proof: The result follows by induction on m.

It is easy to verify that (23) holds when $m = 1, 2$. Suppose $E\left\{\lambda^n\right\} < C^n n^{n-2}$, for $n = 1, 2, ..., m$. Use the following recursive formula [19] to evaluate $E\left\{\lambda^{m+1}\right\}$, which is given by

$$E\left\{\lambda^{m+1}\right\} = \sum_{k=1}^{m+1} \beta' \sum_{m_1+\ldots+m_k=m+1} E\left\{\lambda^{m_1-1}\right\} \cdots E\left\{\lambda^{m_k-1}\right\}.$$

Then we have

$$
\begin{aligned}
E\left\{\lambda^{m+1}\right\} &= \beta'\left(1+mE\{\lambda\}+E\{\lambda^m\}+\sum_{k=2}^{m-1}\sum_{m_1+\ldots+m_k=m+1} E\left\{\lambda^{m_1-1}\right\}\ldots E\left\{\lambda^{m_k-1}\right\}\right)\\
&< \beta'\left(1+m\beta'+C^m m^{m-2}+\sum_{k=2}^{m-1}\sum_{m_1+\ldots+m_k=m+1}\prod_{i=1}^{k} C^{m_i-1}m_i^{m_i-3}\right)\\
&< \beta'\left(1+m\beta'+C^m m^{m-2}+\sum_{k=2}^{m-1}\sum_{m_1+\ldots+m_k=m+1} C^{m+1-k}m^{m-1-k}\right)\\
&< C^{m+1}\left(1+m^{m-1}+\sum_{k=2}^{m-1}\begin{pmatrix} m \\ k-1\end{pmatrix} m^{m-1-k}\right)\\
&< C^{m+1}\left(1+m^{m-1}+\sum_{k=1}^{m-2}\begin{pmatrix} m-1 \\ k\end{pmatrix} m^{m-1-k}\right)\\
&= C^{m+1}\sum_{k=0}^{m-1}\begin{pmatrix} m-1 \\ k\end{pmatrix} m^{m-1-k}\\
&= C^{m+1}(1+m)^{m-1},
\end{aligned}
$$

where the first inequality is based the assumption on $n = 1, \ldots, m$ and the fact that $E\{\lambda\} = \beta'$; the third inequality applies the condition that $C > \max(1, \beta')$ and $m^{m-1} > m^{m-2} + m$ for $m > 2$. This concludes the proof. ■

Applying Stirling's formula and Lemmas I.1,2,3, we can obtain the conclusion that the eigenvalue distribution of $\mathbf{SS}^T$ in the shifted model is identical to that of the independent model, thus assuring the assumption that the columns of $\mathbf{S}$ can be regarded as independent in the large system limit.

APPENDIX II
PROOF OF LEMMA III.2

Proof: From the definition of $\delta\mathbf{a}_f$, we have

$$
E\{\delta\mathbf{a}_f\} = -\frac{1}{M}\left(E\{\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\} - E\{\delta\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\}\right). \tag{24}
$$

We consider the term $E\{\delta\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\}$ first. It is easy to check that (recall that $\mathbf{s}_{kl}$ denotes the spreading code of user k along path l)

$$
\frac{1}{M}E\left\{\left(\delta\mathbf{S}^T\delta\mathbf{S}\right)_{ij}\right\} = \frac{1}{M}\sum_{m=1}^{M}\mathbf{s}_{pq}^T(m)\mathbf{s}_{rs}(m)E\left\{\delta b_p\delta b_r\right\}
$$

$$= \begin{cases} 0, & \text{if } p \neq r \\ \frac{4P_e}{M} \sum_{m=1}^{M} \mathbf{s}_{pq}^T(m)\mathbf{s}_{rs}(m), & \text{if } p = r \end{cases},$$

where $p = \left\lceil \frac{i}{L} \right\rceil$, $q = \mathrm{mod}\,(i, L)$, $r = \left\lceil \frac{j}{L} \right\rceil$, $s = \mathrm{mod}\,(j, L)$. It should be noted we applied the fact that $E\left\{\delta b_p \delta b_r\right\} = 4P_e$ in the second equality.

According to Assumption II.3, the spread codes are mutually independent for different users or different paths. Thus, by applying the strong law of large numbers, we have

$$\frac{1}{M} \sum_{m=1}^{M} \mathbf{s}_{pq}^T(m)\mathbf{s}_{rs}(m) \rightarrow \begin{cases} 0, & \text{if } (p,q) \neq (r,s) \\ 1, & \text{if } (p,q) = (r,s) \end{cases}.$$

Therefore, we have

$$\frac{1}{M} E\left\{\left(\delta\mathbf{S}^T\delta\mathbf{S}\right)_{ij}\right\} \rightarrow \begin{cases} 0, & \text{if } i \neq j \\ \frac{4P_e}{M}, & \text{if } i = j \end{cases}, \qquad \text{almost surely, as } M \rightarrow \infty$$

Similarly, we can show that

$$\frac{1}{M} E\left\{\left(\mathbf{S}^T\delta\mathbf{S}\right)_{ij}\right\} \rightarrow \begin{cases} 0, & \text{if } i \neq j \\ \frac{2P_e}{M}, & \text{if } i = j \end{cases}, \qquad \text{almost surely, as } M \rightarrow \infty$$

This completes the proof. ■

APPENDIX III
PROOF OF PROP. III.3

Proof: The covariance matrix $\mathbf{\Sigma}_f$ is given by

$$\begin{aligned}
\mathbf{\Sigma_f} &\triangleq \frac{1}{M^2}\mathrm{cov}\left(\hat{\mathbf{S}}^T\delta\mathbf{S}\mathbf{a}\right) \\
&= \frac{1}{M^2}E\left\{\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\mathbf{S}\right\} - \frac{1}{M^2}E\left\{\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\delta\mathbf{S}\right\} \\
&- \frac{1}{M^2}E\left\{\delta\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\mathbf{S}^T\right\} + \frac{1}{M^2}E\left\{\delta\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\delta\mathbf{S}\right\} \\
&- E\{\delta\mathbf{a}_f\}E\{\delta\mathbf{a}_f\}^H. \qquad (25)
\end{aligned}$$

The elements in $\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\mathbf{S}$ are given by

$$\left(\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\mathbf{S}\right)_{ij} = \sum_{p=1}^{M}\sum_{q=1}^{M}\sum_{k=1}^{KL}\sum_{l=1}^{KL} \tilde{\mathbf{s}}_i^T(p)\delta\tilde{\mathbf{s}}_k(p)\tilde{\mathbf{s}}_j^T(q)\delta\tilde{\mathbf{s}}_l(q)\mathrm{a}_k\mathrm{a}_l^*,$$

where $\tilde{\mathrm{s}}_i(p) \triangleq b_{\lceil \frac{i}{L} \rceil}(p)\mathrm{s}_{\lceil \frac{i}{L} \rceil,\,\mathrm{mod}(i,L)}(p)$, namely the spreading code (incorporating the channel symbol) of the $\mathrm{mod}(i,L)$-th path of user $\lceil \frac{i}{L} \rceil$ at symbol period p, $\delta\tilde{\mathrm{s}}_i(p) \triangleq \delta b_{\lceil \frac{i}{L} \rceil}(p)\mathrm{s}_{\lceil \frac{i}{L} \rceil,\,\mathrm{mod}(i,L)}(p)$ and a_k is the k-th element of vector $\mathbf{a}$ and equals $a_{\lceil \frac{k}{L} \rceil, mod(k,L)}$. To compute the corresponding expectation, we apply the following properties, which are based on Assumption II.4:

- When $p = q$, if $\lceil \frac{k}{L} \rceil = \lceil \frac{l}{L} \rceil$, $P(\delta\tilde{\mathrm{s}}_k(p) \neq 0, \delta\tilde{\mathrm{s}}_l(q) \neq 0) = P_e$, since $\delta\tilde{\mathrm{s}}_k(p)$ and $\delta\tilde{\mathrm{s}}_l(p)$ are determined by the same decision feedback;
- When $p = q$, if $\lceil \frac{k}{L} \rceil \neq \lceil \frac{l}{L} \rceil$, $P(\delta\tilde{\mathrm{s}}_k(p) \neq 0, \delta\tilde{\mathrm{s}}_l(q) \neq 0) = P_e^2$, since $\delta\tilde{\mathrm{s}}_k(p)$ and $\delta\tilde{\mathrm{s}}_l(p)$ are determined by decision feedback from different users;
- When $p \neq q$, $P(\delta\tilde{\mathrm{s}}_k(p) \neq 0, \delta\tilde{\mathrm{s}}_l(q) \neq 0) = P_e^2$, since $\delta\tilde{\mathrm{s}}_k(p)$ and $\delta\tilde{\mathrm{s}}_l(p)$ are determined by decision feedback from different symbol periods;
- When $\delta\tilde{\mathrm{s}}_k(p) \neq 0$, $\delta\tilde{\mathrm{s}}_k(p) = 2\tilde{\mathrm{s}}_k(p)$.

Thus the expectation of $i-j$th element of $\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\mathbf{S}$ is given by

$$\begin{aligned}
& E\left\{\left(\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\mathbf{S}\right)_{ij}\right\} \\
= \; & 4P_e \sum_{p=1}^{M}\sum_{k=1}^{KL}\sum_{\lceil \frac{l}{L} \rceil = \lceil \frac{k}{L} \rceil} \tilde{\mathrm{s}}_i^T(p)\tilde{\mathrm{s}}_k(p)\tilde{\mathrm{s}}_j^T(p)\tilde{\mathrm{s}}_l(p)\mathrm{a}_k\mathrm{a}_l^* \\
+ \; & 4P_e^2 \sum_{p=1}^{M}\sum_{k=1}^{KL}\sum_{\lceil \frac{l}{L} \rceil \neq \lceil \frac{k}{L} \rceil} \tilde{\mathrm{s}}_i^T(p)\tilde{\mathrm{s}}_k(p)\tilde{\mathrm{s}}_j^T(p)\tilde{\mathrm{s}}_l(p)\mathrm{a}_k\mathrm{a}_l^* \\
+ \; & 4P_e^2 \sum_{\substack{p,\,q\,=\,1 \\ p\,\neq\,q}}^{M}\sum_{k=1}^{KL}\sum_{l=1}^{KL} \tilde{\mathrm{s}}_i^T(p)\tilde{\mathrm{s}}_k(p)\tilde{\mathrm{s}}_j^T(q)\tilde{\mathrm{s}}_l(q)\mathrm{a}_k\mathrm{a}_l^* \\
= \; & T_1 + T_2 + T_3,
\end{aligned}$$

where T_1, T_2 and T_3 represent the corresponding three summations, respectively.

Applying the strong law of large numbers and the assumption on the spreading codes that $\{\tilde{\mathrm{s}}_i(p)\}$ are independent for different values of i or p, we can obtain that, as $M \to \infty$, the following conclusions hold almost surely:

$$\frac{1}{M}T_1 \;\to\; \begin{cases} 4P_e\left(|\mathrm{a}_i|^2 + \frac{1}{N}\sum_{k=1,\,k\neq i}^{KL}|\mathrm{a}_k|^2\right), & \text{if } i = j, \\ 4P_e\left(1+\frac{1}{N}\right)\mathrm{a}_i\mathrm{a}_j^*, & \text{if } i \neq j \text{ and } \lceil \frac{i}{L} \rceil = \lceil \frac{j}{L} \rceil, \\ 0, & \text{if } \lceil \frac{i}{L} \rceil \neq \lceil \frac{j}{L} \rceil \end{cases}$$

$$\frac{1}{M}T_2 \rightarrow \begin{cases} 4P_e^2\left(1+\frac{1}{N}\right)\mathrm{a}_i\mathrm{a}_j^*, & \text{if } \lceil\frac{i}{L}\rceil \neq \lceil\frac{j}{L}\rceil, \\ 0, \quad \text{if } \lceil\frac{i}{L}\rceil = \lceil\frac{j}{L}\rceil \end{cases}$$

$$\frac{1}{M^2}T_3 \rightarrow 4P_e^2\mathrm{a}_i\mathrm{a}_j^*, \qquad \forall i,j.$$

We can apply the same manipulation and obtain that $E\left\{\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\delta\mathbf{S}\right\} = E\left\{\delta\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\mathbf{S}\right\} = \frac{1}{2}E\left\{\delta\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\delta\mathbf{S}\right\}$ as $M\rightarrow\infty$. Therefore, we can obtain (7) since the sum of the middle three terms in (25) is zero and T_3 cancels $E\{\delta\mathbf{a}_f\}E\{\delta\mathbf{a}_f\}^H$. ■

It should be noted that the above analysis is also valid for asynchronous case when P_e is sufficiently small. Similar to the discussion in Section III.A, we can remove part of the chips in the first and the last symbol periods to obtain a similar matrix $\mathbf{S}_{NM-d_{max}\times KL}$, where d_{max} denotes the largest time offsets of different users, measured in chips. When P_e is sufficiently small and M is sufficiently large, we can ignore the terms scaled by P_e^2 and the edge effect in the first and last symbol period. Then, we have

$$E\left\{\left(\mathbf{S}^T\delta\mathbf{S}\mathbf{a}\mathbf{a}^H\delta\mathbf{S}^T\mathbf{S}\right)_{ij}\right\} \approx 4P_e\sum_{k=1}^{KL}\sum_{\lceil\frac{l}{L}\rceil=\lceil\frac{k}{L}\rceil}\tilde{\mathbf{s}}_i^T\tilde{\mathbf{s}}_k\tilde{\mathbf{s}}_j^T\tilde{\mathbf{s}}_l\mathrm{a}_k\mathrm{a}_l^*,$$

where $\tilde{\mathbf{s}}_k$ is the k-th column of matrix $\mathbf{S}$, which converges to T_1 as $M\rightarrow\infty$.

REFERENCES

[1] P. Alexander, A. Grant and M. C. Reed, "Iterative detection of code-division multiple-acess with error control coding," *European Trans. Telecommun.*, Vol. 9, pp. 419–426, Aug. 1998.

[2] P. Alexander and A. Grant, "Iterative channel and information sequence estimation in CDMA," *Proceedings of IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications*, pp. 593–597, Parsippany, NJ, Sept. 2000.

[3] Z. D. Bai, J. W. Silverstein and Y. Q. Yin, "A note on the largest eigenvalue of a large dimensional sample covariance matrix ," *Journal Multivariate Anal.*, Vol. 26, pp. 166–168, 1998.

[4] L. R. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, Vol. 20, pp. 284–287, Aug. 1974.

[5] P. Billingsley, *Probability and Measure*. John Wiley and Sons Inc, New York, US, 1995.

[6] S. Buzzi and H. V. Poor, "Channel estimation and multiuser detection in long-code DS/CDMA systems," *IEEE J. Select. Areas Commun.*, Vol. 19, pp. 1476–1487, Aug. 2001.

[7] S. Buzzi, M. Lops and S. Sardellitti, "Performance of iterative data detection and channel estimation for single-antenna and multiple-antenna wireless communications," *Proceedings of 2003 Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 2003.

[8] S. Buzzi, and H. V. Poor, "A multi-pass approach to joint data and channel estimation in long-code CDMA systems," *IEEE Trans. Wireless Commun.*, Vol. 3, pp. 612–626, March, 2004.

[9] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inform. Theory*, Vol. 46, pp. 2059–2078, Aug. 2000.

[10] G. H. Golub and C. F. Van Loan, *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, 1983.

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, Cambridge, UK, 1985.

[12] J. Hou, J. E. Smee, H. D. Pfister and S. Tomasin, "Implementing interference cancellation to increase the EV-DO Rev A reverse link capacity," *IEEE Commun. Mag.*, Vol. 44, pp. 96–102, Feb. 2006.

[13] M. Kobayashi, J. Boutros and G. Caire, "Successive interference cancellation with SISO decoding and EM channel estimation," *IEEE J. Select. Areas Commun.*, Vol. 19, pp. 1450–1460, Aug. 2001.

[14] E. Kreyszig, *Introductory Functional Analysis With Applications*. John Wiley and Sons Inc, New York, 1989.

[15] A. Lampe, "Iterative multiuser detection with integrated channel estimation for coded DS-CDMA," *IEEE Trans. Commun.*, Vol. 50, no. 8, pp. 1217–1223, Aug. 2002.

[16] C. Laot, A. Glavieux and J. Labat, "Turbo equalization: Adaptive equalization and channel decoding jointly optimized," *IEEE J. Select. Areas Commun.*, Vol. 19, pp. 1744–1752, Sept. 2001.

[17] H. Li and H. V. Poor, "Performance of channel estimation in long code DS-CDMA with and without decision feedback," *Proceedings of the 2003 Conference on Information Sciences and Systems*, The Johns Hopkins University, Baltimore, MD, March 2003.

[18] H. Li and H. V. Poor, "Impact of channel estimation error on multiuser detection via the replica method," *EURASIP Journal on Wireless Communicaitons and Networking*, Vol.2005, pp. 175–186, May 2005.

[19] L. Li, A. M. Tulino and S. Verdú, "Asymptotic eigenvalue moments for linear multiuser detection," *Communications in Information and Systems*, Vol.1, no. 3, pp. 273 – 304, Sept. 2001.

[20] M. Loncar, R. Müller, J. Wehinger, C. Mecklenbraeuker, and T. Abe, "Iterative channel estimation and data detection in frequency-selective fading MIMO channels," *European Transactions on Telecommunications*, Vol.15, no.53, pp.459–470, Sept./Oct. 2004.

[21] R. Lupas and S. Verdú, "Linear multiuser detectors for synthronous code-division multiple-acess channels," *IEEE Trans. Inform. Theory*, Vol. 35, pp. 123–136, Aug. 1989.

[22] M. Moher, "An iterative multiuser decoder for near capacity communications," *IEEE Trans. Commun.*, Vol. 46, no.7, pp. 870–880, July 1998.

[23] H. Niu and J. A. Ritcey, "Iterative channel estimation and decoding of pilot symbol assisted LDPC coded QAM over flat fading channels," *Proceedings of 2003 Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 2003,

[24] H. V. Poor, "Iterative multiuser detection," *IEEE Signal Processing Magazine*, Vol. 21, No,1, pp. 81–86, 2004.

[25] R. Raheli, A. Polydoros, and C. Tzou, "Per-survivor processing: A general approach to MLSE in uncertain environments," *IEEE Trans. Commun.*, Vol. 43, pp. 354–364, Feb./March/April, 1995.

[26] M. C. Reed, C. B. Schlegel, P. D. Alexander and J. A. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: Near single user performance," *IEEE Trans. Commun.*, Vol. 46, no.12, pp. 1693–1699, Dec. 1998.

[27] M. C. Valenti and B. D. Woerner, "Iterative multiuser detection for convolutionally coded asynchronous DS-CDMA," in *IEEE Int. Symp. Personal, Indoor, Mobile Radio Commun.*, Boston, Sept. 1998.

[28] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, Vol. 32, pp. 85–96, Jan. 1986.

[29] S. Verdú, *Multiuser Detection*. Cambridge University Press, Cambridge, UK, 1998.

[30] X. Wang and H. V. Poor, *Wireless Communication Systems: Advanced Techniques for Signal Reception*. Prentice-Hall, Upper Saddle River, NJ, 2004.

[31] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, Vol. 47, no.7, pp. 1046–1061, July 1999.

[32] X. Wang and H. V. Poor, "Subspace methods for blind channel estimation and multiuser detection in CDMA systems," *Wireless Networks*, Vol. 6, pp. 59–71, Feb. 2000.

[33] Z. Xu and M. K.Tsatsanis, “Blind channel estimation for long code multiuser CDMA systems,” *IEEE Trans. Signal Processing*, Vol. 48, pp. 988–1001, Apr. 2000.

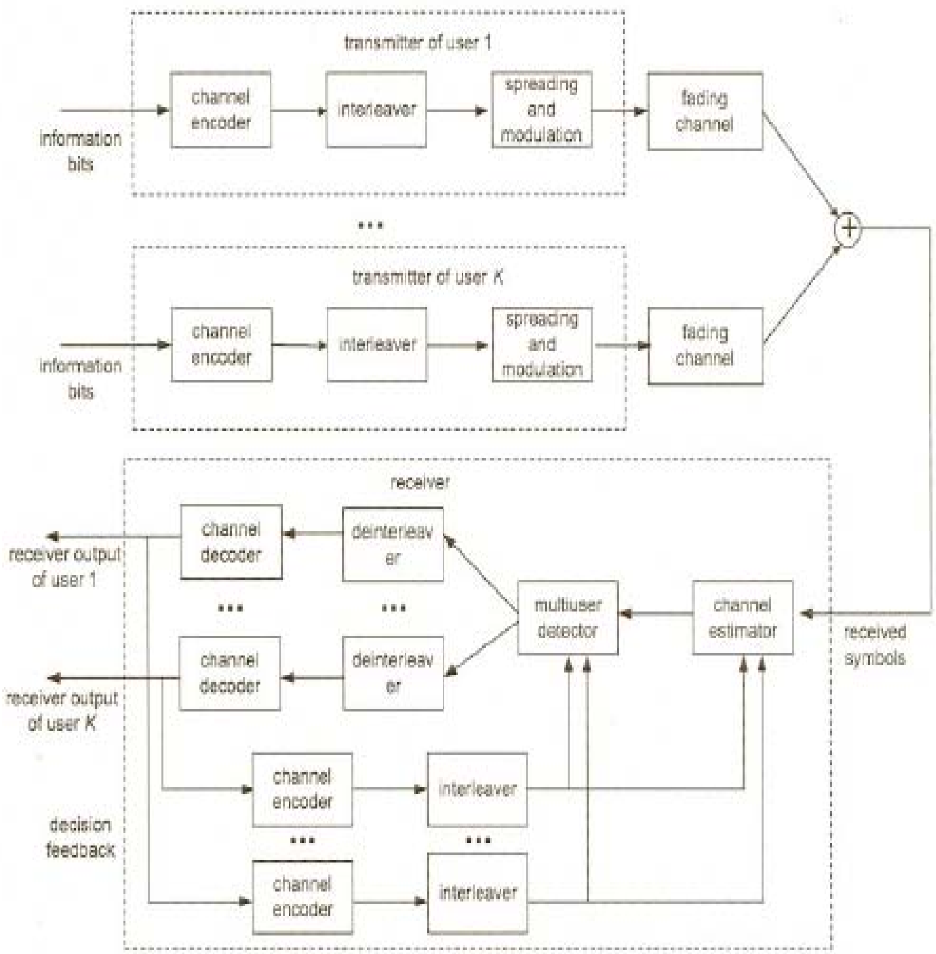

Fig 1. CDMA system with an iterative receiver

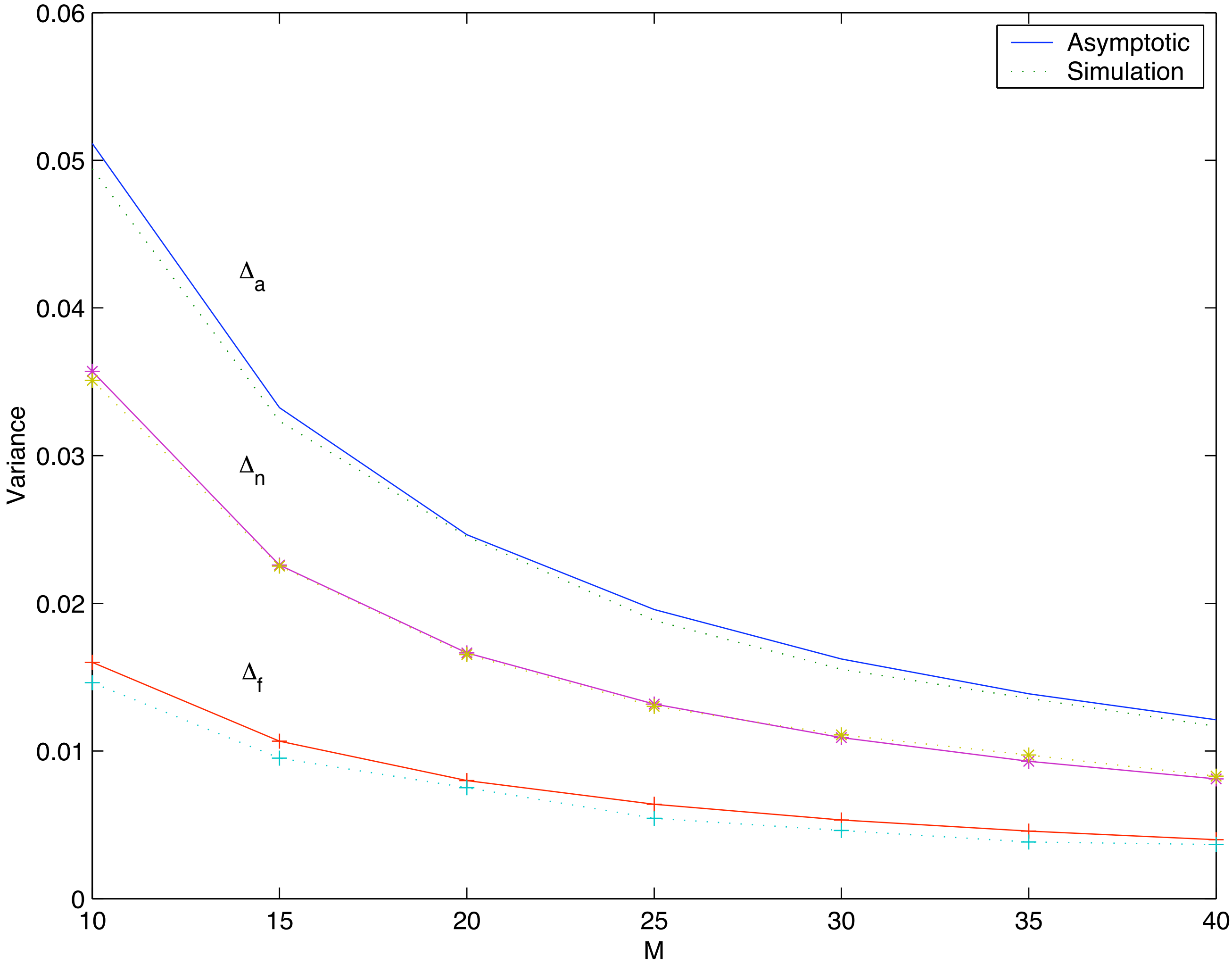

Fig. 2. Average variance of channel estimates versus the coherence time M

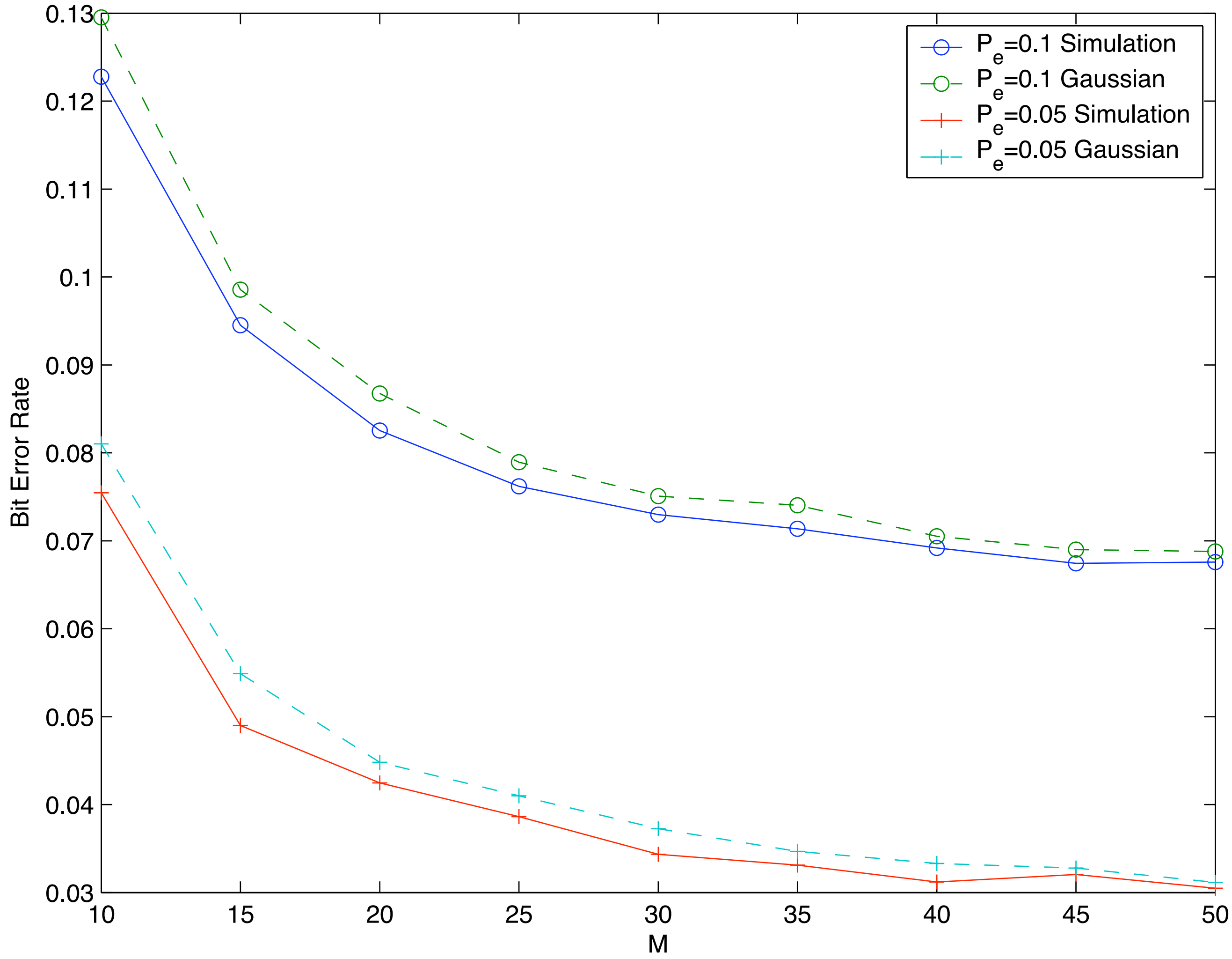

Fig. 3. Comparison of simulated bit error rates and those obtained using a Gaussian approximation

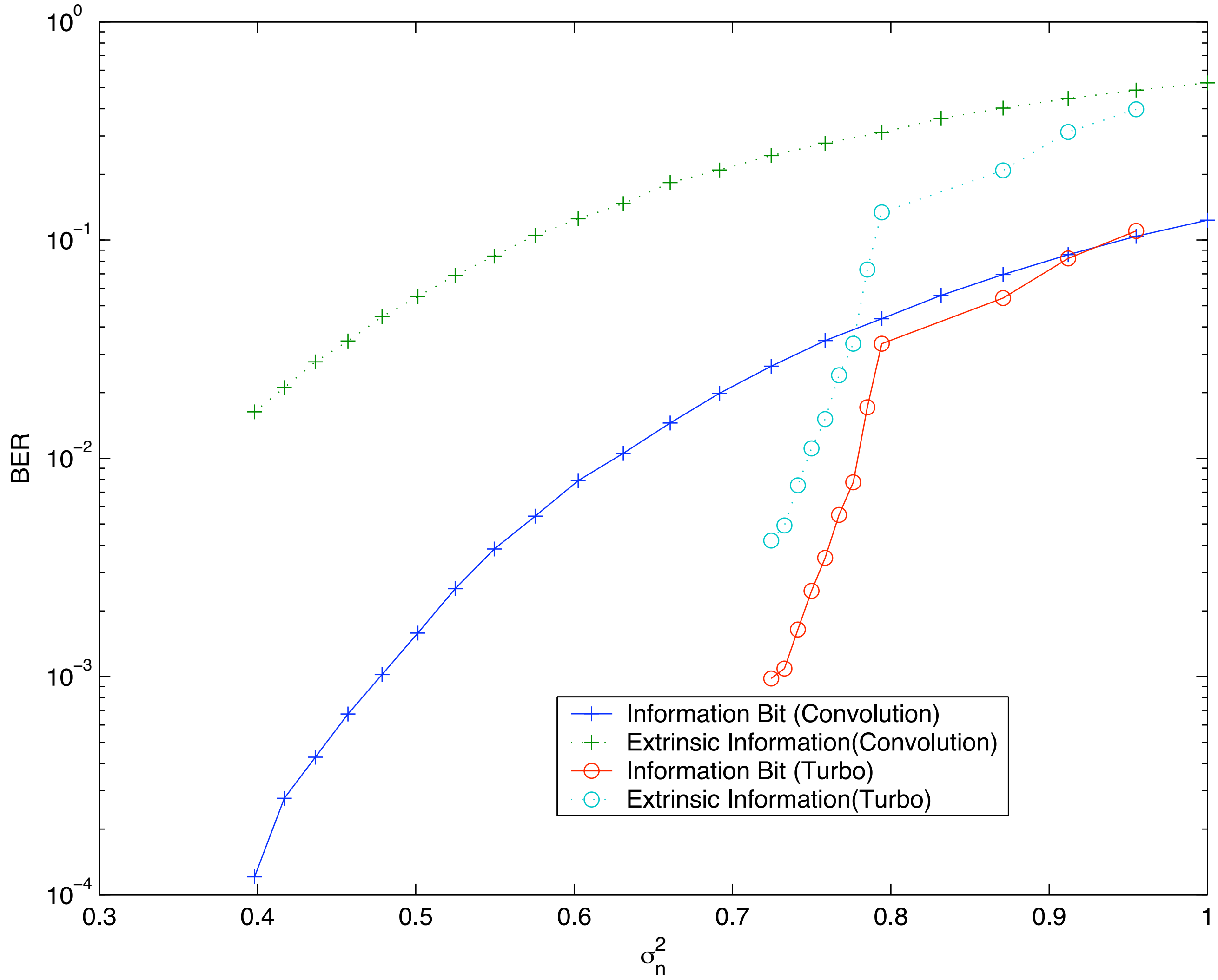

Fig. 4. Performance of channel codes used in the numerical results, where the input SNR = $\frac{1}{\sigma_n^2}$

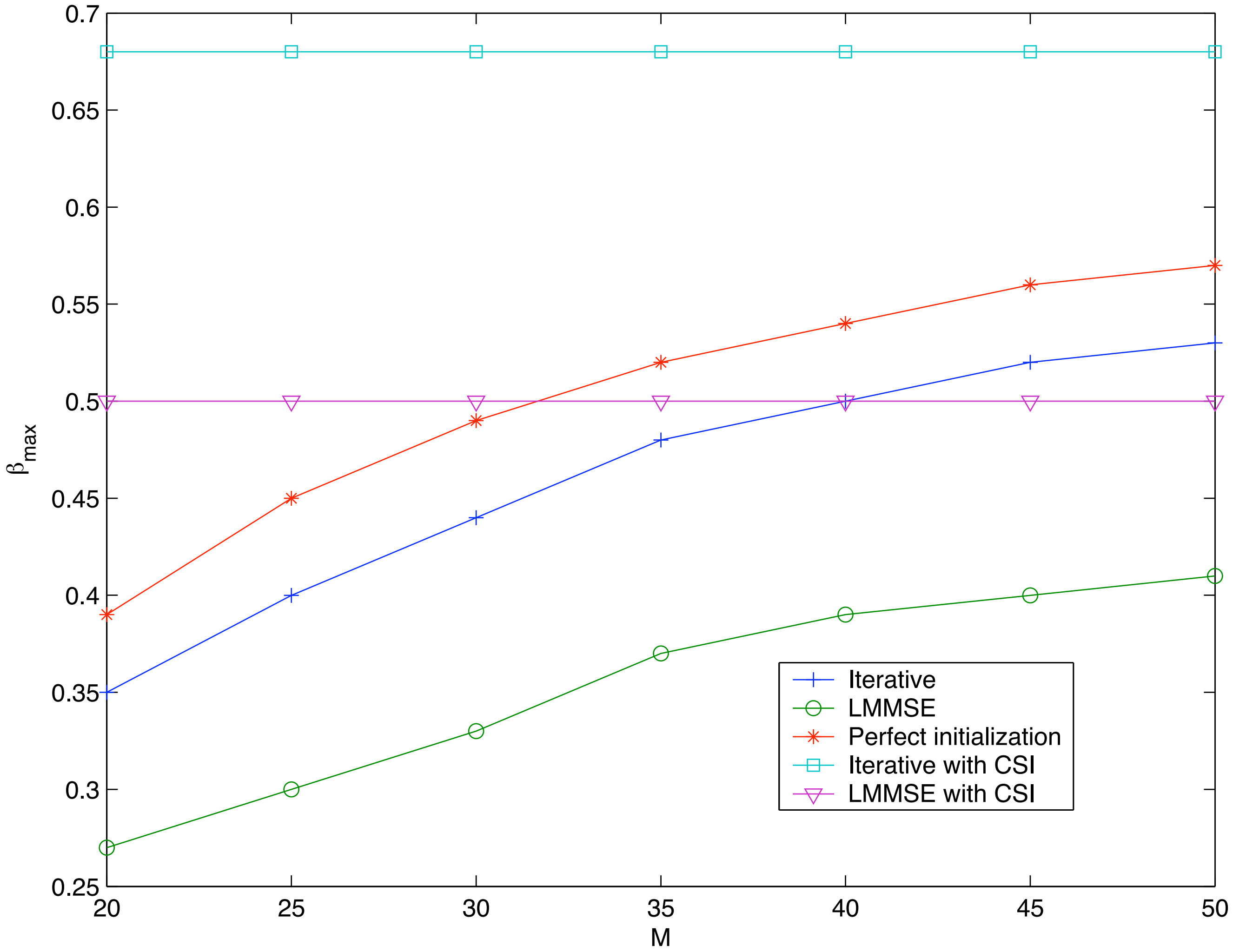

Fig. 5. Maximum load of systems with convolutional codes

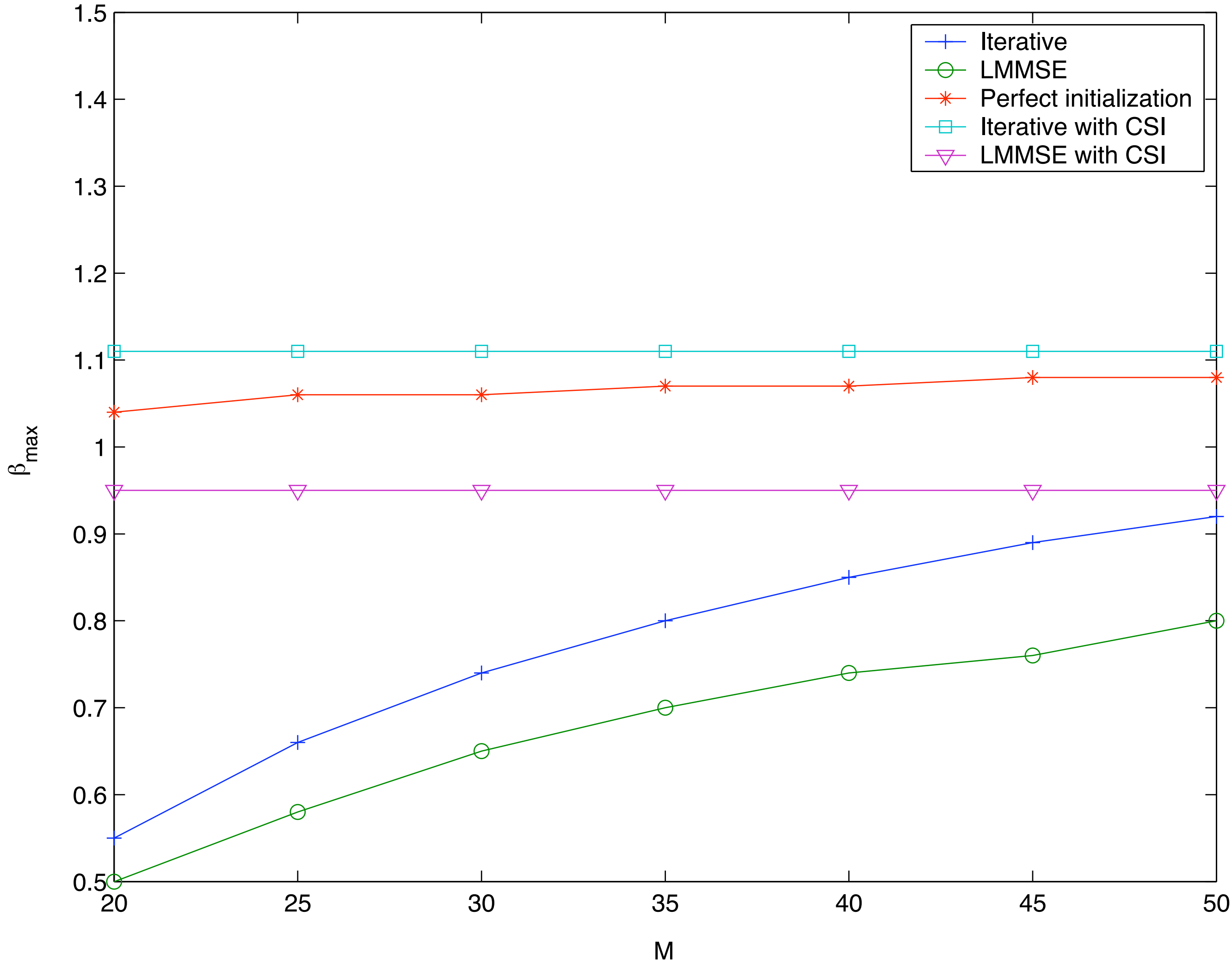

Fig. 6. Maximum load of systems with turbo codes

Appendix B

Energy Efficiency in Multi-hop CDMA Networks: A Game Theoretic Analysis*

Sharon Betz and H. Vincent Poor
Department of Electrical Engineering
Princeton University
{sbetz,poor}@princeton.edu

Abstract

A game-theoretic analysis is used to study the effects of receiver choice on the energy efficiency of multi-hop networks in which the nodes communicate using Direct-Sequence Code Division Multiple Access (DS-CDMA). A Nash equilibrium of the game in which the network nodes can choose their receivers as well as their transmit powers to maximize the total number of bits they transmit per unit of energy is derived. The energy efficiencies resulting from the use of different linear multiuser receivers in this context are compared, looking at both the non-cooperative game and the Pareto optimal solution. For analytical ease, particular attention is paid to asymptotically large networks. Significant gains in energy efficiency are observed when multiuser receivers, particularly the linear minimum mean-square error (MMSE) receiver, are used instead of conventional matched filter receivers.

1 Introduction

In a wireless multi-hop network, nodes communicate by passing messages for one another; permitting multi-hop communications, rather than requiring one-hop communications, can increase network capacity and allow for a more ad hoc (and thus scalable) system (with little or no centralized control). For these reasons, and because of their potential for commercial, military, and civil applications, wireless multi-hop networks have attracted considerable attention over the past few years. In these networks, energy efficient communication is important because the nodes are typically battery-powered and therefore energy-limited. Work on energy-efficient communication in these multi-hop networks has often focused on routing protocols; this work instead looks at power control and receiver design choices that can be implemented independently of (and thus in conjunction with) the routing protocol.

One approach that has been very successful in researching energy efficient communications in both cellular and multi-hop networks is the game-theoretic approach described in [1, 2]. Much of the game-theoretic research in multi-hop networks has focused on pricing schemes (e.g. [3, 4]). In this work, we avoid the need for such a pricing scheme by using instead a nodal utility function to capture the energy costs. It further differs from previous research by considering receiver design, as [5] does for cellular networks.

We propose a distributed noncooperative game in which the nodes can choose their transmit power and linear receiver design to maximize the number of bits that they can send per unit of power. After describing the network and internodal communications in Section 2, we derive the Nash equilibrium for this game, as well as for a set of games with set receivers, in Section 3. We then extend the asymptotic work of Tse and Hanly [6] to fit the multihop network structure in Section 4; we apply this in Section 5 to find the Pareto optimal solution in an asymptotically large, SINR-balanced network. Finally we present some numerical results and a conclusion in Sections 6 and 7.

2 System Model

Consider a wireless multi-hop network with K nodes (users) and an established logical topology, where a sequence of connected link-nodes $l \in L(k)$ forms a route originating from a source k (with $k \in L(k)$ by definition). Let $m(k)$ be the node after node k in the route for node k. Assume that all routes that go through a node k continue through $m(k)$ so that node k transmits only to $m(k)$. Nodes communicate with each other using DS-CDMA with processing gain N (N chips per bit).

The signal received at a node m (after chip-matched filtering) sampled at the chip rate over one symbol duration

*This research was supported in part by the U. S. Air Force Research Laboratory and in part by the Defense Advanced Research Projects Agency.

can be expressed as

$$\mathbf{r}^{(m)} = \sum_{k=1}^{K} \sqrt{p_k} h_k^{(m)} b_k \mathbf{s}_k + \mathbf{w}^{(m)} \tag{1}$$

where p_k, b_k, and $\mathbf{s}_k$ are the transmit power, transmitted symbol, and (binary) spreading sequence for node k; $h_k^{(m)}$ is the channel gain between nodes k and m; and $\mathbf{w}^{(m)}$ is the noise vector which is assumed to be Gaussian with mean $\mathbf{0}$ and covariance $\sigma^2\,\mathbf{I}$. (We assume here $p_m = 0$.) Assume the spreading sequences are random, i.e., $\mathbf{s}_k = \frac{1}{\sqrt{N}}[v_1 \ldots v_N]^T$, where the v_i's are independent and identically distributed (i.i.d.) random variables taking values $\{-1, +1\}$ with equal probabilities. Denote the cross-correlations between spreading sequences as

$$\rho_{kj} = \mathbf{s}_k^T \mathbf{s}_j, \tag{2}$$

noting that $\rho_{kk} = 1$ for all k.

Let us represent the linear receiver at the mth node for the kth signature sequence by a coefficient vector $\mathbf{c}_k^{(m)}$. The output of this receiver can be written as

$$\begin{aligned} y &= \mathbf{c}_k{}^T \mathbf{r}^{(m)} &(3)\\ &= \sqrt{p_k} h_k^{(m)} b_k \mathbf{c}_k{}^T \mathbf{s}_k + \sum_{j \neq k} \sqrt{p_j} h_j^{(m)} b_j \mathbf{c}_k{}^T \mathbf{s}_j + \mathbf{c}_k{}^T \mathbf{w}^{(m)}. &(4)\end{aligned}$$

The signal-to-interference-plus-noise ratio (SINR), γ_k, of the kth user at the output of receiver $m(k)$ is

$$\gamma_k = \frac{p_k h_k^{(m(k))^2} \left(\mathbf{c}_k{}^T \mathbf{s}_k\right)^2}{\sigma^2 \mathbf{c}_k{}^T \mathbf{c}_k + \sum_{j\neq k} p_j h_j^{(m(k))^2} \left(\mathbf{c}_k{}^T \mathbf{s}_j\right)^2}. \tag{5}$$

Each user has a utility function that is the ratio of its effective throughput to its transmit power, i.e.,

$$u_k = \frac{T_k}{p_k}. \tag{6}$$

Here, the throughput, T_k, is the net number of information bits sent by k (generated by k or any node whose route goes through k) and received without error at the intended destination, $m(k)$, per unit of time. (We assume that all the congestion control is done in the choice of routing.)

Following the discussion in [5], we will use

$$T_k = \frac{L}{M} R f(\gamma_k) \tag{7}$$

where L and M are the number of information bits and the total number of bits in a packet, respectively (without loss of generality assumed here to be the same for all users); R is the transmission rate, which is the ratio of the bandwidth to the processing gain and is taken for now to be equal for all users; and $f(\cdot)$ is an efficiency function that closely approximates the packet success rate. This efficiency function can be any increasing, continuously differentiable, sigmoidal[1] function with $f(0) = 0$ and $f(+\infty) = 1$. See [5] for more discussion of the efficiency function.

Using (7), (6) becomes

$$u_k = \frac{L}{M} R \frac{f(\gamma_k)}{p_k}. \tag{8}$$

When the receiver used is a matched filter (MF) (i.e. $\mathbf{c}_k^{(m(k))} = \mathbf{s}_k$), the received SINR is

$$\begin{aligned} \gamma_k^{\mathrm{MF}} &= \frac{p_k h_k^{m(k)^2} \left(\mathbf{s}_k{}^T \mathbf{s}_k\right)^2}{\sigma^2 \mathbf{s}_k{}^T \mathbf{s}_k + \sum_{j\neq k} p_j h_j^{m(k)^2} \left(\mathbf{s}_k{}^T \mathbf{s}_j\right)^2} &(9)\\ &= \frac{p_k h_k^{m(k)^2}}{\sigma^2 + \sum_{j\neq k} p_j h_j^{m(k)^2} \rho_{kj}^2}. &(10)\end{aligned}$$

When the receiver is a linear minimum mean-squared error (MMSE) receiver , the filter coefficients and the received SINR are [7]

$$\mathbf{c}_k^{\mathrm{MMSE}} = \frac{\sqrt{p_k} h_k^{m(k)}}{1 + p_k h_k^{m(k)^2} \mathbf{s}_k^T \mathbf{A}_k^{-1} \mathbf{s}_k} \mathbf{A}_k^{-1} \mathbf{s}_k \tag{11}$$

and

$$\gamma_k^{\mathrm{MMSE}} = p_k h_k^{m(k)^2} \mathbf{s}_k^T \mathbf{A}_k^{-1} \mathbf{s}_k, \tag{12}$$

where

$$\mathbf{A}_k = \sigma^2\,\mathbf{I} + \sum_{j\neq k} p_j h_j^{m(k)^2} \mathbf{s}_j \mathbf{s}_j^T. \tag{13}$$

When the receiver is a decorrelator[2] (DE) (i.e. $\mathbf{C} = [\mathbf{c}_1 \cdots \mathbf{c}_K] = \mathbf{S}(\mathbf{S}^T\mathbf{S})^{-1}$ where $\mathbf{S} = [\mathbf{s}_1 \cdots \mathbf{s}_K]$), the received SINR is

$$\gamma_k^{\mathrm{DE}} = \frac{p_k h_k^{m(k)^2}}{\sigma^2 \mathbf{c}_k^T \mathbf{c}_k}. \tag{14}$$

For any linear receiver with all nodes' coefficients chosen independently of their transmit powers (including the MF and DE), as well as for the MMSE receiver,

$$\frac{\partial \gamma_k}{\partial p_k} = \frac{\gamma_k}{p_k}. \tag{15}$$

[1] A continuous increasing function is sigmoidal if there is a point above which the function is concave and below which the function is convex.

[2] Here, we must assume that $K \leq N$.

3 The Noncooperative Power-Control Game

Let $\mathcal{G} = \left[\mathcal{K}, \{A_k\}, \{u_k\}\right]$ denote the noncooperative game where $\mathcal{K} = \{1, \ldots, K\}$ and $A_k = [0, P_{\max}] \times \mathbb{R}^N$ is the strategy set for the kth user. Here, $P_{\max}$ is the maximum allowed power for transmission. Each strategy in A_k can be written as $\mathbf{a}_k = (p_k, \mathbf{c}_k)$, where p_k and $\mathbf{c}_k$ are the transmit power and the receiver filter coefficients, respectively, of user k. Then the resulting noncooperative game can be expressed as the maximization problem for $k = 1, \ldots, K$:

$$\max_{\mathbf{a}_k} u_k = \frac{LR}{M} \max_{p_k, \mathbf{c}_k} \frac{f(\gamma_k(p_k, \mathbf{c}_k))}{p_k}, \tag{16}$$

where γ_k is expressed explicitly as a function of p_k and $\mathbf{c}_k$.

This is similar to the noncooperative power-control game in [5]; here, however, the channel gains are between pairs of nodes rather than between a node and the base-station.

Since the choice of receiver is independent of the transmit power and $f(\cdot)$ is an increasing function, the analysis of [5] applies, so the maximization from (16) becomes:

$$\max_{p_k, \mathbf{c}_k} \frac{f(\gamma_k(p_k, \mathbf{c}_k))}{p_k} = \max_{p_k} \frac{f(\max_{\mathbf{c}_k} \gamma_k(p_k, \mathbf{c}_k))}{p_k}. \tag{17}$$

Note that the MMSE receiver achieves the maximum SINR amongst all linear receivers, so that if a Nash equilibrium exists, at that equilibrium all receivers must be MMSE receivers. Then the maximization problem becomes

$$\max_{p_k} \frac{f(\gamma_k^{\mathrm{MMSE}}(p_k))}{p_k}. \tag{18}$$

Let $\mathcal{G}_{\mathbf{C}} = \left[\mathcal{K}, \{[0, P_{\max}]\}, \{u_k\}\right]$ denote the noncooperative game that differs from $\mathcal{G}$ in that users cannot choose their linear receivers but are forced to use the receive filter coefficients $[\mathbf{c}_1 \cdots \mathbf{c}_K] = \mathbf{C}$ (which may be a function of the powers, $\mathbf{P}$). The resulting noncooperative game can be expressed as the following maximization problem for $k = 1, \ldots, K$:

$$\max_{\mathbf{a}_k} u_k = \max_{p_k} u_k(p_k, \mathbf{c}_k) = \frac{LR}{M} \max_{p_k} \frac{f(\gamma_k^{\mathbf{c}_k}(p_k))}{p_k} \tag{19}$$

where $\gamma_k^{\mathbf{c}_k}$ is expressed explicitly as a function of p_k. Then the maximization problem in (18) is one of the games $\mathcal{G}_{\mathbf{C}}$ when $\mathbf{C}$ is chosen to be the MMSE receivers.

For any $\mathbf{C}$ matrix (or $\mathbf{C}(\mathbf{P})$ for which (15) holds), the utility function for each user is maximized when

$$p_k = \min\{P_k, p_k^*\} \tag{20}$$

where p_k^* is the unique positive number that satisfies

$$f(\gamma_k^{\mathbf{c}_k}(p_k^*)) = \gamma_k^{\mathbf{c}_k}(p_k^*) f'(\gamma_k^{\mathbf{c}_k}(p_k^*)). \tag{21}$$

As long as the users all have the same efficiency function,

$$\gamma_1^{\mathbf{c}_1}(p_1^*) = \ldots = \gamma_K^{\mathbf{c}_K}(p_K^*) = \gamma^* \tag{22}$$

where γ^* is the unique positive number that satisfies

$$f(\gamma^*) = \gamma^* f'(\gamma^*). \tag{23}$$

Finally, since $\frac{f(\gamma_k)}{p_k}$ is quasi-concave[3] in p_k, we can use the result cited in [2, Appendix I]: $\mathcal{G}_{\mathbf{C}}$ has a Nash equilibrium and, as is the case in [5], it is unique. At this equilibrium, unless there is a node k with $p_k^* > P_k$, the powers are such that the nodes are SINR-balanced (i.e. (22) holds).

Returning to the game $\mathcal{G}$, a similar result holds: there exists a unique equilibrium where all receivers are MMSE detectors and, if the power limit is high enough, the powers are SINR-balanced.

4 Asymptotically Large Systems: Extending the Tse-Hanly Equations to Multi-Hop Networks

Assume that the channel gains are independent. That is, in the asymptotic regime when $N, K \to \infty$ while $K/N = \beta$, the interferers' channel gains, $h_k^{(m)^2}$ for all $m \neq k, m(k)$, are iid realizations of the random variable G with pdf f_G, and the primary channel gains, $h_k^{(m(k))^2}$ for all k, are iid realizations of the random variable H with pdf f_H (where $f_H(h) = 0 \forall h \leq 0$). Let $q = \mathbb{P}\{m(j) = m(k)\}$ for all $j \neq k$.

We can apply results from [6] to analyze the nodes' SINRs. Then we find a probability density function for p such that in an asymptotically large system where all nodes have powers distributed by this function, with probability one all nodes have SINR of at least γ for some γ. If this distribution is not unique, we choose the one that minimizes the nodes' powers. For simplicity, and since we are considering the asymptotic regime, we assume that the distribution of p_k is independent of all channel gains except for $h_k^{(m(k))^2}$. For convenience of notation, let $f_{p,H}(\cdot,\cdot) = f_{p_k, h_k^{(m(k))^2}}(\cdot,\cdot)$, and note $\int_0^\infty f(p,h)dp = f_H(h)$ for all h. Then the joint density of p_k, $h_k^{(m(k))^2}$, and $h_k^{(m(j))^2}$ for $j \neq k$ is

$$\begin{aligned} f_{p_k, h_k^{(m(k))^2}, h_k^{(m(j))^2}}(p,h,g) &= f_{p,H}(p,h)\delta(g-h)q \\ &\quad + f_{p,H}(p,h) f_G(g)(1-q). \end{aligned} \tag{24}$$

Applying the results from [6], when the receiver at node k is a matched filter, decorrelator, or MMSE receiver, the random SINR at the receiver converges in probability as

[3] A function is quasi-concave if there exists a point below which the function is nondecreasing and above which the function is non-increasing.

$N, K \to \infty$ while $K/N = \beta$. These asymptotic SINRs are uniquely described by the equations (where $j \neq k$):

$$\gamma^{\mathrm{MF}} = \frac{p_k h_k^{(m(k))^2}}{\sigma^2 + \beta \mathbb{E}\left[p_j h_j^{(m(k))^2}\right]} \tag{25}$$

$$\gamma^{\mathrm{DE}} = \begin{cases} \frac{p_k h_k^{(m(k))^2}(1-\beta)}{\sigma^2}, & \alpha < 1; \\ 0, & \alpha \geq 1. \end{cases} \tag{26}$$

and

$$\gamma^{\mathrm{MMSE}} = \frac{p_k h_k^{(m(k))^2}}{\sigma^2 + \beta \int_0^\infty dp \int_0^\infty dg f_P(p) f_G(g) I(pg, p_k h_k^{(m(k))^2}, \gamma^{\mathrm{MMSE}})}, \tag{27}$$

where $I(a,b,c) = \frac{ab}{b+ac}$.

If the nodes choose their transmit powers so that the SINRs are balanced, the following theorem determines what SINRs are achievable at all receivers as well as the minimum transmit powers to achieve any achievable SINR when the nodes use the MMSE receiver, under the assumptions listed above.

Theorem 4.1. *A necessary and sufficient condition for an SINR, γ, to be achievable is for*

$$\beta\gamma q \frac{1}{1+\gamma} + \beta\gamma(1-q)\mathbb{E}\left[\frac{\mathsf{G}}{\mathsf{H}+\gamma\mathsf{G}}\right] < 1. \tag{28}$$

When (28) holds, each user can achieve the desired SINR, γ, and the minimum power solution to do so is to assign each node, k, transmit power

$$p_k = P_{\mathrm{MMSE}}\left(h_k^{(m(k))^2}, \gamma\right) \tag{29}$$

$$= \frac{1}{h_k^{(m(k))^2}} \cdot \frac{\gamma\sigma^2}{1 - \beta\gamma q \frac{1}{1+\gamma} - \beta\gamma(1-q)\mathbb{E}\left[\frac{\mathsf{G}}{\mathsf{H}+\gamma\mathsf{G}}\right]}. \tag{30}$$

4.1 Proof of Theorem 4.1

We start with a lemma that is a straightforward consequence of the definition of $I(a,b,c)$.

Lemma 4.2. *For all positive real numbers a_0, a, b, c, $a_0 \leq a$ if and only if $I(a_0, b, c) \leq I(a, b, c)$.*

Then the proof follows.

Proof. To show necessity, assume that there is a pdf f with $\int_0^\infty f(p,h)dp = f_H(h)$ for all h, such that in an asymptotically large system where all nodes have powers and primary channel gains distributed by f, with probability one all nodes have SINR when using an MMSE receiver of at least γ for some set γ. Let $Q = \inf\{ph : f(p,h) > 0\}$. Then

$$\begin{aligned}
\frac{Q}{\gamma} &\geq \sigma^2 + \beta \int_0^\infty dg \int_0^\infty dp \int_0^\infty dh f_{p_k, h_k^{(m(k))^2}, h_k^{(m(j))^2}}(p,h,g) I(pg, Q, \gamma) \\
&= \sigma^2 + \beta q \int_0^\infty dp \int_0^\infty dh f_{p,H}(p,h) I(ph, Q, \gamma) \\
&+ \beta(1-q) \int_0^\infty dg \int_0^\infty dp \int_0^\infty dh f_{p,H}(p,h) f_G(g) I(ph\frac{g}{h}, Q, \gamma) \\
&\geq \sigma^2 + \beta q \int_0^\infty dp \int_0^\infty dh f_{p,H}(p,h) I(Q, Q, \gamma) \\
&+ \beta(1-q) \int_0^\infty dg \int_0^\infty dp \int_0^\infty dh f_{p,H}(p,h) f_G(g) I(Q\frac{g}{h}, Q, \gamma) \\
&= \sigma^2 + \beta q \frac{Q}{1+\gamma} + \beta(1-q) \int_0^\infty dg \int_0^\infty dh f_H(h) f_G(g) \frac{gQ}{h+\gamma g} \\
&= \sigma^2 + Q\beta q \frac{1}{1+\gamma} + Q\beta(1-q)\mathbb{E}\left[\frac{\mathsf{G}}{\mathsf{H}+\gamma\mathsf{G}}\right].
\end{aligned} \tag{31}$$

This implies that

$$Q\left(1 - \beta\gamma q \frac{1}{1+\gamma} - \beta\gamma(1-q)\mathbb{E}\left[\frac{\mathsf{G}}{\mathsf{H}+\gamma\mathsf{G}}\right]\right) \geq \gamma\sigma^2 > 0, \tag{32}$$

so $\beta\gamma q \frac{1}{1+\gamma} + \beta\gamma(1-q)\mathbb{E}\left[\frac{\mathsf{G}}{\mathsf{H}+\gamma\mathsf{G}}\right] < 1$, proving necessity.

When (28) holds, it is easy to show that $P_{\mathrm{MMSE}}(h, \gamma)$ is positive for all primary channel gains, h. It is also straightforward to show that if each node, k, uses transmit power $P_{\mathrm{MF}}\left(h_k^{(m(k))^2}, \gamma\right)$, all nodes will achieve the SINR requirement, γ, finishing the proof of sufficiency.

Finally, consider any other joint distribution of powers and primary channel gains whose marginal distribution for H is f_H, and let Q^* be the minimal received power in this distribution. Then by exactly the same argument as was used in the proof of necessity,

$$Q^* \geq \frac{\gamma\sigma^2}{1 - \beta\gamma q \frac{1}{1+\gamma} - \beta\gamma(1-q)\mathbb{E}\left[\frac{\mathsf{G}}{\mathsf{H}+\gamma\mathsf{G}}\right]} \tag{33}$$

$$= h P_{\mathrm{MMSE}}(h, \gamma), \ \forall h > 0. \tag{34}$$

This means that assigning powers according to P_{MMSE} does indeed give the minimal power solution. □

5 A Global Optimization Problem

A useful global optimization problem is

$$\max \sum_{k=1}^{K} \alpha_k u_k = \frac{L}{M} R \max \sum_{k=1}^{K} \frac{\alpha_k f(\gamma_k)}{p_k}, \tag{35}$$

where the α_k's are set weighting variables. This problem is equivalent to finding a Pareto-optimal solution of the game.

According to [5], even in the special case of a cellular system where $L(k) = \{k\}$ for all nodes $k = 1, 2, \ldots, K$ and all nodes are transmitting to the base-station, "Pareto-optimal solutions are, in general, difficult to obtain." For simplicity, we restrict the problem by requiring that the solution is "fair": all nodes have equal receiver output SINRs (i.e. SINR-balancing), so $\gamma = \gamma_1 = \gamma_2 = \ldots \gamma_K$.

With this assumption, (35) becomes

$$\frac{L}{M} R \ \max f(\gamma) \sum_{k=1}^{K} \frac{\alpha_k}{p_k}. \tag{36}$$

For the matched filter, we can apply (5) with $m = m(k)$ to see that the users' SINRs are equal if and only if

$$\left(B + \left(\frac{1}{\gamma} + 1 \right) D \right) \mathbf{p}(\gamma) = \sigma^2 \mathbf{1} \tag{37}$$

where B is a K by K matrix with entries $B_{kj} = -h_j^{(m(k))^2} \rho_{kj}^2$, D is K by K diagonal matrix with diagonal entries $D_{kk} = h_k^{(m(k))^2}$, and $\mathbf{1}$ is a vector of K ones.

The SINR that maximizes (36) is the γ that satisfies

$$0 = \frac{\partial}{\partial \gamma} \left[f(\gamma) \sum_{k=1}^{K} \frac{\alpha_k}{p_k(\gamma)} \right] \tag{38}$$

$$= \frac{\partial}{\partial \gamma} \left[f(\gamma) \right] \sum_{k=1}^{K} \frac{\alpha_k}{p_k(\gamma)} - f(\gamma) \sum_{k=1}^{K} \frac{\alpha_k}{p_k^2(\gamma)} \frac{\partial}{\partial \gamma} \left[p_k(\gamma) \right], \tag{39}$$

where $p_k(\gamma)$ and $\frac{\partial}{\partial \gamma} [p_k(\gamma)]$ are the kth elements of

$$\mathbf{p}(\gamma) = \sigma^2 \left(B + \left(\frac{1}{\gamma} + 1 \right) D \right)^{-1} \mathbf{1} \tag{40}$$

and

$$\frac{\partial}{\partial \gamma} [\mathbf{p}(\gamma)] = \sigma^2 \left(\gamma B + (1 + \gamma) D \right)^{-1} D \left(\gamma B + (1 + \gamma) D \right)^{-1} \mathbf{1}. \tag{41}$$

For the decorrelator, it is easy to show that the non-cooperative results are equal to the globally optimal results, since the users' achieved SINRs are independent of all the powers of all interferers.

Finally, for the MMSE receiver, we can apply the results from Section 4. In a large system, if all users choose their transmit powers based on the values of $h_k^{(m(j))}$ for $m(j) \neq m(k)$ only through the average of these interference gains and if we use the assumptions of Section 4, the SINR is approximated by

$$\gamma_k^{\mathrm{MMSE}} \simeq \frac{p_k h_k^{(m(k))^2}}{\sigma^2 + \frac{1}{N} \sum_{j \neq k} I(p_j h_j^{(m(k))^2}, p_k h_k^{(m(k))^2}, \gamma_k^{\mathrm{MMSE}})}. \tag{42}$$

Any γ_k which satisfies $\frac{\partial \gamma_k}{\partial p_k} = \frac{\gamma_k}{p_k}$ is a solution to (42).

Then the power for user k to achieve the SINR γ^* is

$$p_k^{\mathrm{MMSE}} = \frac{1}{h_k^{(m(k))^2}} \frac{\gamma^* \sigma^2}{1 - \beta \gamma^* \left(q \frac{1}{1+\gamma^*} + (1 - q) \zeta(\gamma^*) \right)}, \tag{43}$$

where $\zeta(\gamma)$ is the mean value of $\frac{\mathsf{G}}{\mathsf{H}+\gamma \mathsf{G}}$ in the network. Equal received SINRs amongst the users is achieved with minimum power consumption when $p_k h_k^{(m(k))^2} = \kappa(\gamma)$ is constant for all k and

$$\kappa(\gamma) = \frac{\gamma \sigma^2}{1 - \beta \gamma \left(q \frac{1}{1+\gamma} + (1 - q) \zeta(\gamma) \right)}. \tag{44}$$

Then, (36) can be expressed as

$$\frac{L}{M} R \left(\sum_{k=1}^{K} \alpha_k h_k^{(m(k))^2} \right) \max_{\gamma} \frac{f(\gamma)}{\kappa(\gamma)}. \tag{45}$$

The solution to $\max_\gamma \frac{f(\gamma)}{\kappa(\gamma)}$ must satisfy $\frac{\partial}{\partial \gamma} \left(\frac{f(\gamma)}{\kappa(\gamma)} \right) = 0$. Combining this with (44) gives the equation that must be satisfied by the solution to the maximization problem in (45):

$$f(\gamma) = \gamma f'(\gamma) \left(1 - \frac{\frac{\beta q \gamma}{(1+\gamma)^2} + \beta(1 - q) \gamma \zeta(\gamma)}{1 - \frac{\beta q \gamma^2}{(1+\gamma)^2} - \beta(1 - q) \gamma \zeta(\gamma)} \right). \tag{46}$$

If $\zeta(\gamma) << 1$, then the equation is approximately the same as in the cellular case [5] with $K/N \to \beta q$. Then, the ability to use multiple hops to communicate, and therefore reduce transmit power, has similar results to reducing the system load; furthermore, for a large range of values of βq, the MMSE target SINRs for the noncooperative game and for the Pareto-optimal solution are close.

6 Numerical Results

Consider a multi-hop network with $K = 100$ nodes distributed randomly in a square 500 meters by 500 meters surrounding an access point in the center. We use a simple routing scheme where all nodes transmit to the closest node that is closer to the access point (or the access point if that is closest). We assume that each packet contains 100 bits of data and no overhead ($L = M = 100$); the transmission rate is $R = 100$ kb/s; the thermal noise power is $\sigma^2 = 5 \times 10^{-16}$ Watts; the channel gains are distributed with a Rayleigh distribution with mean $0.3 d^{-2}$, where d is the distance between the transmitter and receiver; and the processing gain is N. We use the same efficiency function as [5], namely $f(\gamma) = (1 - e^{-\gamma})^M$.

Table 1 shows the average utility for four representative sets of randomly chosen spreading sequences, one for each of N = 50, 100, 200, and 300, comparing the mean utility under the various power choice method discussed above. Table 2 shows the target SINRs for the socially optimal results displayed in Table 1.

	MF	DE	MMSE
$N = 50$			
non-coop.	0		1.198×10^{10}
soc. opt.	2.025×10^{-14}		1.199×10^{10}
$N = 100$			
non-coop.	0	1.095×10^{8}	1.417×10^{10}
soc. opt.	1.050×10^{-4}	1.095×10^{8}	1.417×10^{10}
$N = 200$			
non-coop.	0	7.459×10^{9}	1.476×10^{10}
soc. opt.	2.512×10^{-10}	7.459×10^{9}	1.476×10^{10}
$N = 300$			
non-coop.	0.2056	1.001×10^{10}	1.493×10^{10}
soc. opt.	1.351×10^{9}	1.001×10^{10}	1.493×10^{10}

Table 1. Mean utilities for four representative sets of spreading sequences.

N	MF	DE	MMSE
50	0.87		6.39
100	1.31	6.47	6.43
200	0.99	6.47	6.45
300	5.03	6.47	6.46

Table 2. Socially optimal SINRs for the same four representative sets of spreading sequences.

The socially optimally implemented MF receiver performs poorly in heavily-loaded systems, while the non-cooperative implementation fails to achieve non-zero utility except in the case with the lightest load. Even in the case where $\beta = 1/3$, the mean utility for the socially optimal MF receiver is less than a tenth of the MMSE receiver's mean utility. Using the DE receiver (for which we require that $K \geq N$), as was noted in Section 5, there is no difference between the non-cooperative and socially optimal results: both cases have the same target SINR and thus the same mean utility. For the MMSE receiver, this difference between the mean utility in the non-cooperative and socially optimal implementations is very small. Finally, the DE and MMSE receivers both significantly outperform the MF receiver in all four of these cases. There is, however, a price to pay in using the better-performing receivers: these receivers require more information at every node as well as significantly more computation. These issues will be addressed further in later research.

7 Conclusion

We have analyzed the cross-layer issue of energy-efficient communication in multi-hop networks using a game theoretic method. Focusing on linear receivers, we have derived the transmit power levels that results in a Nash equilibrium for multiple receiver designs, showing that at this equilibrium the users are SINR-balanced. We then generalized the important asymptotic work of Tse and Hanly to allow for the case where users and their interferers may be transmitting to different locations, keeping the cellular example as a special case. We applied these asymptotic results, as well as exact results for the MF and DE receivers, to find the equations for the SINR-balanced Pareto-optimal solution. We showed that the MMSE receiver is the optimal receiver and that in many cases the non-cooperative MMSE receiver results are quite close to the socially optimal results.

References

[1] D. J. Goodman and N. B. Mandayam, "Power control for wireless data," *IEEE Personal Communications*, vol. 7, no. 2, pp. 48–54, April 2000.

[2] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 291–303, February 2002.

[3] O. Ileri, S.-C. Mau, and N. B. Mandayam, "Pricing for enabling forwarding in self-configuring ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 151–162, January 2005.

[4] P. Marbach and Y. Qiu, "Cooperation in wireless ad hoc networks: A market-based approach," *IEEE/ACM Transactions on Networking*, vol. 13, no. 6, pp. 1325–1338, December 2005.

[5] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, "An energy-efficient approach to power control and receiver design in wireless data networks," *IEEE Transactions on Communications*, vol. 53, no. 11, pp. 1885–1894, November 2005.

[6] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 641–657, March 1999.

[7] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge University Press, 1998.